

Model studies of the structures, reactivities, and reaction mechanisms of metalloenzymes

by K. Morokuma
D. G. Musaev
T. Vreven
H. Basch
M. Torrent
D. V. Khoroshun

Electronic structure theory, which in recent years has been actively and effectively applied to the modeling of chemical reactions involving transition-metal complexes, is now also being applied to the modeling of biological processes involving metalloenzymes. In the first part of this paper, we review our recent electronic structure studies using relatively simple models of two metalloenzymes—methane monooxygenase and ribonucleotide reductase. In the second part of the paper, we review a new hybrid theoretical method we have developed for modeling the reactivities of large molecular systems. We describe the limitations of these models and indicate how they may be further improved to reliably model the reactivities of complicated metalloenzymes.

1. Introduction

Modeling the reactivities of biological systems is currently one of the long-range goals of computational chemistry. In the field of bioinorganic chemistry, which includes metalloenzyme chemistry, because of improved

spectroscopic techniques, the elucidation of the structure of active sites of many enzymes and the nature of the key intermediates in enzyme-catalyzed processes, and an ever-increasing crystallographic database of protein structures, sufficient material has been accumulated to allow the study of many enzymatic catalytic processes in detail at the molecular level [1].

In the last few decades, quantum-mechanical calculations of many chemical reactions including organic, inorganic, and organometallic reactions have provided new qualitative and quantitative insights that have not been available from experimental studies alone. For instance, it was rare ten years ago that organometallic chemists envisioned the structure and energetics of transition states in proposing a reaction mechanism. Today, they often visualize the transition-state structure and collaborate with quantum chemists to find such transition states computationally in order to substantiate a proposed mechanism. The Nobel Prize in Chemistry awarded to John A. Pople and Walter Kohn in 1998 represents the recognition that electronic structure calculation is making a unique and essential contribution to chemistry.

Computational modeling of the structure and functions of metalloenzymes has been slow, because 1) not enough details of relevant reaction mechanisms have been known

©Copyright 2001 by International Business Machines Corporation. Copying in printed form for private use is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the *Journal* reference and IBM copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free without further permission by computer-based and other information-service systems. Permission to *republish* any other portion of this paper must be obtained from the Editor.

0018-8646/01/\$5.00 © 2001 IBM

experimentally for metalloenzyme systems, 2) the electronic structure of metal-containing systems has not been clearly understood, and 3) computational requirements for the modeling of metal-containing systems are very large. However, with problems 1 and 2 being resolved as mentioned above, and computers and codes becoming faster, this field has gained momentum in the last few years. Using rather simple models of enzymatic systems, quantum-mechanical calculations have already led to new insights regarding the behavior of electrons during metalloenzyme reaction processes. In the next section of this review, we restrict our consideration to one case study we performed, the reactivity of methane monooxygenase (MMO) and ribonucleotide reductase (RNR), and discuss some of our findings from the model calculations.

However, the reality of such modeling is still very much restricted by practical computational limitations. As indicated in the next section, computational approaches require the use of relatively simple models of the complicated enzyme system. This is because the electronic structure calculation required to describe such reaction systems demands a great deal of computer time, and our present computers are not fast enough to allow the use of larger and more realistic models. However, as is discussed in Section 3, it appears that even if the computers were faster, one could not apply the theoretical method used in the current model calculations to larger and more realistic models. To achieve realistic results, a highly accurate method is certainly required for the part of the molecule where the reaction is taking place, but it may be possible to adopt less accurate and therefore less computer-demanding methods for the part of the molecule not directly involved in the reaction. We have recently implemented a hybrid approach in which different methods are combined for computations pertaining to one large molecule or a molecular system. In Section 3, we review this "onion-like" method, which we named ONIOM, which facilitates such computations.

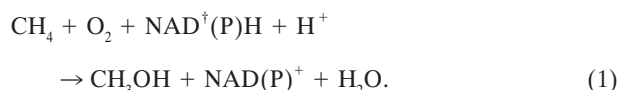
Although the ONIOM method has not yet been used for the study of metalloenzyme systems, in Section 4 we briefly discuss its applicability to more realistic models of enzymatic reactions.

2. Density-functional studies of the reactivities of methane monooxygenase and ribonucleotide reductase

Binuclear Fe-containing metalloenzymes, such as hemerythrin [2], ribonucleotide reductase (RNR) [3], methane monooxygenase (MMO) [3(a, b), 4], Δ -desaturase [5], phenol hydroxylase [6], xylene monooxygenase [7], alkane hydroxylase [7], toluene 2-monooxygenase [8], toluene 4-monooxygenase [9], purple acid phosphatase [10], rubrerythrin [11], and ferritin [12], have been the focus of intensive research for the past

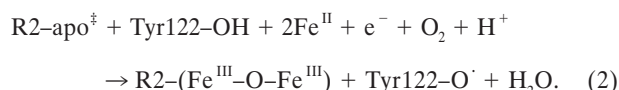
decade or so. Here we discuss the geometric and electronic properties of the active sites of MMO and RNR, the plausible mechanism of dioxygen coordination and activation on the MMO and RNR, and the mechanism of hydrocarbon hydroxylation by MMO.

Methane monooxygenase (MMO) is an enzyme that catalyzes the methane oxidation reaction, i.e., conversion of the inert methane molecule to methanol [1]:



The best-characterized forms of the soluble MMO (sMMO) [1, 13] contain three protein components: hydroxylase (MMOH), the so-called B component (MMOB), and reductase (MMOR), each of which is required for efficient substrate hydroxylation coupled to NADH oxidation. The hydroxylase, MMOH, which binds O_2 and the substrate and catalyzes oxidation, is a dimer, each half of which contains three subunit types (α , β , γ) and a hydroxy-bridged binuclear iron cluster. In the resting state of MMOH (MMOH_{ox}), the diiron cluster is in the diferric state [$\text{Fe}^{\text{III}}\text{--Fe}^{\text{III}}$], and can accept one or two electrons to generate the mixed-valence [$\text{Fe}^{\text{III}}\text{--Fe}^{\text{II}}$] or diferrous [$\text{Fe}^{\text{II}}\text{--Fe}^{\text{II}}$] states, respectively. The diferrous state of hydroxylase (MMOH_{red}) is the only one capable of reacting with dioxygen and initiating the catalytic cycle.

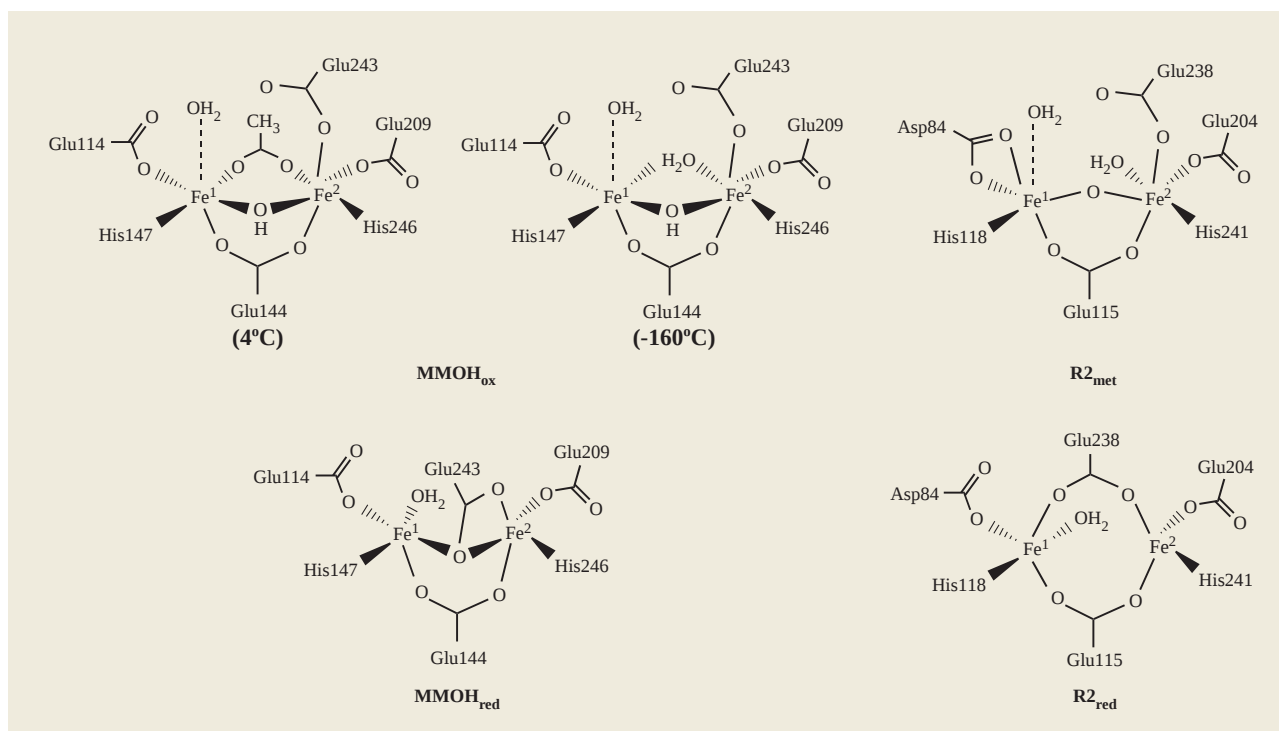
RNR catalyzes the reduction of ribonucleotides and generates deoxyribonucleotides [14]. In the literature there exist at least four classes of RNR, each of which has its unique composition and cofactor requirement [15]. The best-characterized class (I) has an $\alpha_2\beta_2$ structure with α_2 and β_2 dimers, referred to as R1 and R2 units. Each R2 unit contains a binuclear non-heme Fe active site. This class of RNR requires the presence of O_2 to generate the tyrosyl radical (Tyr) and binuclear active site via oxidation of a binuclear ferrous center of R2 to an oxo-bridged binuclear ferric site:



Extensive studies of the core structures of the active sites of MMOH and R2 show some sequence homology between these two enzymes, and a similar flexibility of the ligand environment of the Fe-centers [1, 16]. This ligand flexibility has been postulated to be one of the important factors for the proper functioning of the enzymes [17]. Indeed, X-ray studies of the core structure of oxidized MMOH (MMOH_{ox}) isolated from *M. capsulatus* show that at 4°C the two Fe atoms are triply bridged by one hydroxo and two μ -1,2-carboxylate ligands with an Fe-Fe

[†] Nicotinamide adenine dinucleotide.

[‡] Apo-enzyme.



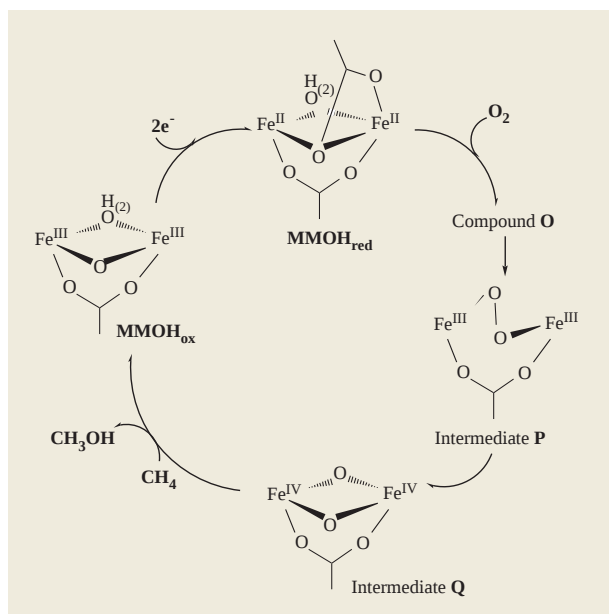
Scheme 1

Structural representation of the binuclear Fe center of MMOH_{ox} , R2_{met} , MMOH_{red} , and R2_{red} (see Reference [1]).

distance of ~ 3.4 Å [18]. Each Fe-center has one histidine ligand. In addition, Fe^1 has one terminal aquo and one carboxylate ligand, while Fe^2 has two carboxylate ligands, as shown in **Scheme 1**. On the other hand, the structure recorded at -160°C shows an aquo bridge replacing one of the carboxylate bridges; the resulting structure has one $\mu\text{-OH}$, one $\mu\text{-H}_2\text{O}$, and one carboxylate bridge with a shorter Fe–Fe distance (~ 3.1 Å) [18(b)]. Similarly, the oxidized binuclear active site of the R2 subunit of RNR from *E. coli* (R2_{met}) shows one $\mu\text{-oxo}$ and one 1,2-carboxylate bridge with an Fe–Fe distance of ~ 3.2 Å, and one histidine ligand for each Fe-center [19]. Fe^1 has a terminal aquo and a chelating carboxylate ligand, whereas Fe^2 has two monodentate terminal carboxylate ligands (as in MMOH), and one terminal aquo ligand.

As shown in the literature [1], the oxidized forms of MMOH and R2 (MMOH_{ox} and R2_{met}), including two ferric atoms, Fe^{III} , are the resting state of these enzymes. Only their two-electron reduced forms, MMOH_{red} and R2_{red} , with two ferrous Fe^{II} iron centers are capable of reacting with O_2 —the reaction that initiates both of the catalytic cycles described above. Structural studies [1, 18–20] demonstrated that two-electron reduction of MMOH_{ox} and R2_{met} dramatically changes the ligand environment of the Fe centers. Indeed, for MMOH,

during reduction two hydroxo/aquo bridging ligands move out, and one of the carboxylate ligands of Fe^2 , shifts to form a monodentate bridge between the two metals and coordinates with the Fe^2 -center in a bidentate manner [18(c)]. Therefore, the two Fe atoms are changed from six-coordinate (6C) to five-coordinate (5C), with one vacant site for each Fe. Spectroscopic studies are in accord with such a (5C, 5C) assignment for the reduced state of MMOH [21]. Similarly, in R2_{met} the two-electron reduction leads to dissociation of two oxo/aquo bridging ligands [20]. However, upon reduction of R2_{met} , the terminal ligand Asp84 shifts from chelating to monodentate terminal bridging to Fe^1 , a shift which has no analog in the counterpart reduction of MMOH. The carboxylate ligand Glu238 in R2, which moves from terminal monodentate in Fe^2 to a bridging position, also gives a bidentate bridge rather than monodentate, as in MMOH. Thus, according to crystallographic studies [20], R2_{red} has two approximately equivalent 4C Fe-centers. However, it is believed that one of the Fe-centers, Fe^1 , also has an aquo ligand (see Scheme 1). This is consistent with 1) spectroscopic data [22] indicating that Fe^1 and Fe^2 atoms in R2_{red} are 5C and 4C, respectively, and 2) the reduced binuclear active site of manganese-substituted *E. coli*, which has been found to have a water ligand bound to Mn [1, 23].



Scheme 2

Experimentally proposed catalytic cycle of MMO.

Thus, the data described above clearly show that the ubiquitous carboxylate ligands (glutamates and aspartates) of MMOH and R2 may play a special role in the proper functioning of these enzymes. Indeed, the flexibility of these ligands allows them, depending on enzymatic needs, to coordinate with the Fe-centers as bidentate ligands (bridging or chelating) when saturation of the first coordination sphere of Fe-centers is required, and as monodentate (terminal) ligands when one or more vacant coordination sites in the Fe-centers are needed to enable a certain reaction step to take place. Therefore, modeling of the flexibility of the MMOH and R2 ligand environment and the factors affecting this flexibility are important tasks. Modeling the flexibility is the first task of our computational studies.

It has been established in the literature [1] that MMOH_{red} reacts very rapidly with O_2 , forming a metastable, so-called compound **O** that converts spontaneously to another compound designated as **P**, (Scheme 2). Spectroscopic studies [24] indicate that compound **P** is a peroxide species, in which both oxygens are bound symmetrically to the iron atoms. The structure of this compound is unknown. Compound **P** converts spontaneously to compound **Q**, which has been proposed to contain two antiferromagnetically coupled high-spin Fe^{IV} centers. EXAFS and spectroscopic studies [25] of compound **Q** have demonstrated that it has a diamond-core, $(\text{Fe}^{\text{IV}})_2(\mu\text{-O})_2$ structure with one short (1.77 Å)

and one long (2.05 Å) Fe–O bond per Fe atom and a short Fe–Fe distance of 2.46 Å. Compound **Q** has been proposed to be the key oxidizing species for MMO [1].

In contrast with MMOH, in R2 the second intermediate, **X**, of the two intermediates **U** and **X** observed for the reaction of O_2 with R2, is found to be a spin-coupled $[\text{Fe}^{\text{III}}\text{Fe}^{\text{IV}}]$ complex containing two or three μ -oxo-bridges, but no peroxo-level intermediate has been observed. The intermediate **U** is believed to be a protonated tryptophan radical [1(b)]. Therefore, the second task of our computational studies is to model the mechanism of the dioxygen coordination and O–O bond activation on MMOH and R2 enzymes.

The third important problem related to MMO and RNR is to elucidate the mechanism of the oxygen-atom incorporation into the substrate from the key species of the reaction of MMOH_{red} and R2_{red} with O_2 , to form compounds **Q** and **X**, respectively. In the literature, several mechanisms [1, 26] have been proposed for the reaction of intermediate **Q** with alkanes; these mechanisms can be divided into two different classes: radical and nonradical. The radical mechanism begins with abstraction of the hydrogen atom from the substrate to form **QH** (hydroxyl-bridged **Q**-compound) and a free alkyl radical, whereas the nonradical mechanism implies a concerted pathway, occurring via a four-center transition state and leading to the “hydrido-alkyl-**Q**” compound. The latest experimental studies [27] show that the mechanism of this reaction is more complex, and could even be different for the enzyme isolated from different organisms. Radical clock substrate probe studies with sMMO isolated from *M. capsulatus* (Bath) afforded no evidence for the formation of a substrate radical, whereas for the enzyme isolated from *M. trichosporium* OB3b, evidence for a substrate radical was detected. Thus, the elucidation of the validity of different mechanisms including the structure of assumed intermediates and transition states is an open question and requires extensive further studies.

To provide some insight into these extremely important problems, we have carried out comprehensive computational modeling studies. However, in order to perform high-level quantum chemical calculations on these systems, several nontrivial questions must first be answered.

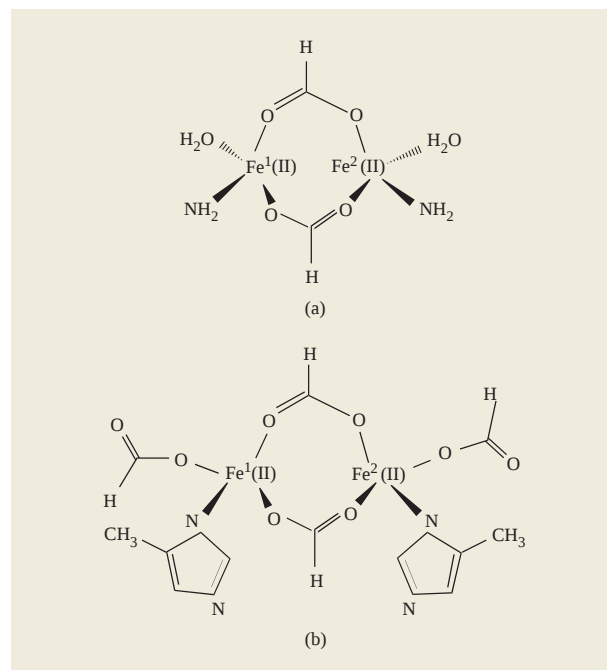
The first question concerns the choice of a reasonable model for the enzymes incorporating all available experimental findings. For example, according to the experimental results and previous modeling work [28–30], a reasonable model of MMOH_{red} and R2_{red} should include two Fe–Fe bridging carboxylate ligands, and one imidazole and one monodentate-coordinated carboxylate ligand for each Fe center. Furthermore, imidazole rings of the His ligands should be located cis to each other. Here, we use two models, shown in Scheme 3. The “small” model

includes two Fe–Fe bridging C₁ carboxylates, one NH₂, and H₂O ligands coordinated with each Fe, which respectively model His and monodentate carboxylate residues. The “medium” model used here replaces NH₂ ligands with an imidazole ring, and H₂O ligands with the C₁ carboxylate ligands in the small model. This model additionally includes a water molecule coordinated with one of the Fe (Fe¹) centers.

Today, these smaller model compounds, in which O–O bond cleavage and oxygen atom incorporation into the substrate take place, can be studied using a high-level computational method. However, one should note that neither model takes into account the protein environment, and, consequently, the existence of different amino acids on the active sites of MMOH and RNR. In other words, in these models, MMOH_{red} and R2_{red} species are different from each other ONLY in the coordination mode of one of the bridging carboxylate ligands; in MMOH_{red} both bridging carboxylate ligands are μ -1,2-type bridges between two Fe centers, while in R2_{red} one of these carboxylates is a μ -1,1-type bridge. Henceforth, therefore, we call the structures with two μ -1,2-type carboxylate bridges “R2-like” structures, while complexes with one μ -1,2-type and one μ -1,1-type bridge are labeled as “MMOH-like” complexes.

The second important question in these studies concerns the spin state of the system. For example, spectroscopic studies [4] indicate that the binuclear iron cluster of MMOH_{ox} (diferric) has two $S = 5/2$ iron atoms antiferromagnetically spin-coupled (AFSC) to yield a diamagnetic center with the exchange coupling constant J between two Fe-centers of $7 \pm 3 \text{ cm}^{-1}$ [4]. However, diferrous MMOH (MMOH_{red}) exhibits ferromagnetic spin coupling (FSC) of two $S = 2$ ferrous iron atoms to give an $S = 4$ ground state [1, 21]. Thus, MMOH_{ox} and many other intermediates (see below) are AFSC diamagnetic species, while MMOH_{red} is an FSC paramagnetic species. In other words, during the hydrocarbon hydroxylation reaction by MMO, one may expect spin-crossing of the system and switching between antiferromagnetic and ferromagnetic couplings among the expected intermediates and transition states. The spectroscopic picture for R2_{red} is also complex. On the basis of Mössbauer studies, the two Fe-centers in R2_{red} are high-spin ferrous ions [31]. EPR studies of R2_{red} show a very weak integer spin signal, considered to derive from a small fraction of molecules having FSC sites [32]. MCD studies show a paramagnetic center with a saturation behavior indicating two Fe-centers with $M_S = \pm 2$ at the ground state. However, a spin Hamiltonian analysis of the saturation magnetization behavior indicates that the two Fe atoms are weakly ($J \sim 0.5 \text{ cm}^{-1}$) AFSC [33].

The third question involves the choice of adequate computational methods and approximations. Here, the



Scheme 3

Models used in the present studies. (a) “Small” model; (b) “medium” model.

problem should be divided into two parts. The first is an intrinsic problem of how to represent the open-shell low-spin coupling ($2S + 1 = 0$) of the two paramagnetic centers using the single-determinant methods (such as DFT); the “restricted” calculations for the low-spin state give the closed-shell singlet state, which cannot represent an open-shell atom. An “unrestricted” calculation with low M_S value gives a heavily spin-contaminated state, or the total wave function becomes symmetry-broken and does not represent a true eigenstate. Here one should use a multireference method (such as MCSCF or CASPT2), which, unfortunately, is impractical for such large systems. Therefore, it is sometimes more practical (when the magnitude of the spin coupling between the two centers is not strong) to ignore the antiferromagnetic nature of these systems, and to perform spin-unrestricted open-shell single-determinant calculations for ferromagnetically coupled high-spin states. This approach also retains the proper spins on individual metal (Fe) atoms. Since the magnitude of the spin coupling between the two centers of MMOH and R2 is very small, as discussed above, we expect that the mechanisms of the reactions studied below are not overly influenced by the antiferromagnetic nature of the complexes.

Using the answers to the three questions as described above, we have carried out the model calculations and

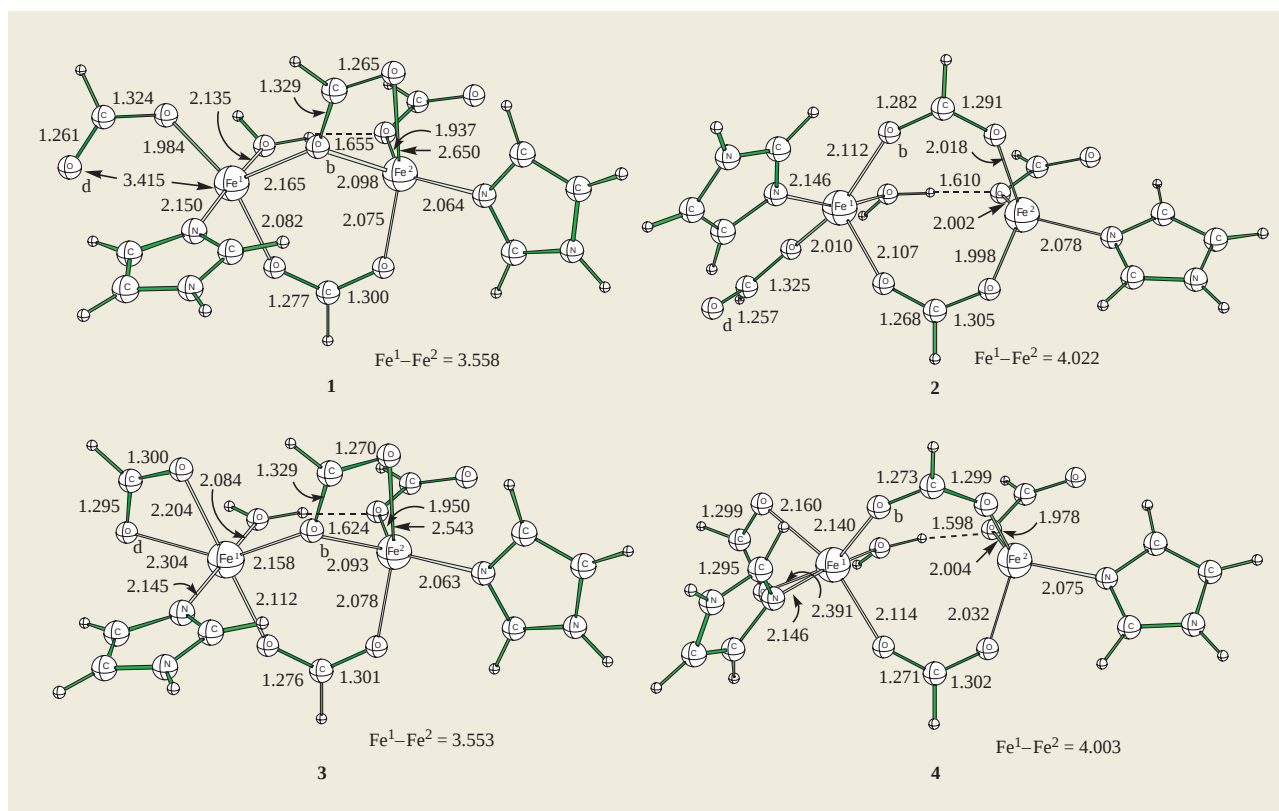


Figure 1

Optimized geometries of the equilibrium structures (distances in Å).

describe the results as follows. In the first subsection which follows, we discuss the mechanisms and factors affecting the flexibility of the ligand environment of MMOH_{red} and R2_{red} . In the next subsection, we discuss the mechanism of dioxygen coordination and activation on MMOH_{red} and R2_{red} . The third subsection is devoted to studies of the reaction mechanism of compound **Q** with the CH_4 molecule, leading to the methanol complex. The last subsection presents a perspective for future studies.

Flexibility of the ligand environment of MMOH and R2

Several attempts have been made to elucidate the factors affecting ligand flexibility and the reactivity of MMOH and R2 . Recently, Nordlund and coworkers [17], using mutation experiments, demonstrated that substrate (azide) coordination with R2_{red} could cause a 1,2-carboxylate shift of one of the bridging carboxylate ligands, Glu238. These authors believe that such a carboxylate shift (of Glu238) could be a key feature in understanding the dioxygen activation mechanism at the Fe-centers. Also, extensive biomimetic studies of carboxylate coordination modes in

diiron complexes, as well as O_2 coordination with these biomimetic complexes by Lippard and coworkers [34] have demonstrated that 1) having bulky steric ligands [such as XDK, *m*-xylylenediamine bis(propyl Kemp's triacid)imide, derivatives] in the bridging carboxylate ligands could facilitate a μ -1,1 $\leftrightarrow$ μ -1,2-carboxylate shift; and 2) the 1,2-carboxylate shift could be a rate-determining step for the entire substrate (O_2) coordination and ligand rearrangement process [34].

In connection with these experiments, we have also attempted to investigate the carboxylate shift between two Fe centers (called the "1,2-carboxylate shift") and the bidentate $\leftrightarrow$ monodentate carboxylate rearrangement within one metal center [29]. In these studies we have used the B3LYP density functional theory (DFT) method [35], which is discussed more fully on pages 381 and 382, and have chosen the "medium" model of MMOH_{red} and R2_{red} , presented in Scheme 3. Since the exchange coupling constant of R2_{red} is extremely small, we have considered both MMOH_{red} and R2_{red} as ferromagnetically coupled high-spin species, and we have studied them at their $2M_s + 1 = 9$ spin states.

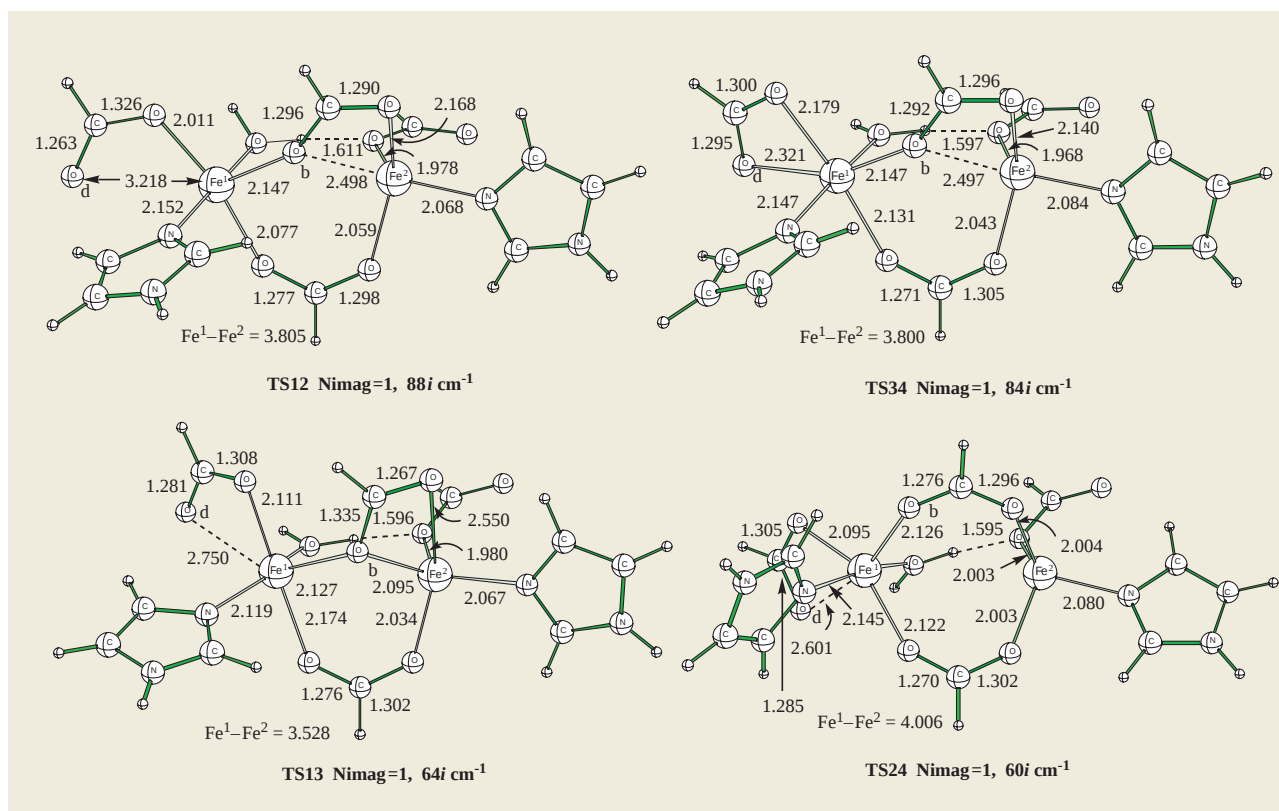


Figure 2

Optimized geometries of the transition states (distances in Å).

Geometries of the calculated equilibrium structures and transition states (TSs) are shown in **Figures 1 and 2**, respectively. In **Figure 3**, we schematically present the calculated mechanisms of the μ -1,1 $\leftrightarrow$ μ -1,2 carboxylate shift, and the bidentate $\leftrightarrow$ monodentate carboxylate rearrangement in MMOH-like and R2-like complexes, together with their corresponding relative energies.

As seen from Figure 1, we have found two sets of minima [29]: a) structures **1** and **3**, both with one monodentate (μ -1,1) and one bidentate (μ -1,2) Fe–Fe bridging carboxylate ligand (MMOH-like structures); and b) structures **2** and **4**, both with two bidentate (μ -1,2) Fe–Fe bridging carboxylate ligands (R2-like structures). Structures **1** and **3**, as well as **2** and **4**, differ from each other only in the coordination mode of the terminal carboxylate ligand (which models Glu114 in MMOH_{red} and Asp84 in R2_{red}) located on Fe¹; the terminal carboxylate ligand binds to the Fe-center monodentately in structures **1** and **2**, but bidentately in structures **3** and **4**. Thus, horizontally (in Figure 3), processes **1** $\leftrightarrow$ **2** and **3** $\leftrightarrow$ **4** involve migration of one of the carboxylate ligands between the two Fe centers, from being a monodentate

bridge (μ -1,1) between the two metals (as well as coordinating with the Fe²-center in a bidentate manner, left), to form a bidentate (μ -1,2) bridge between the two irons (right). This is the so-called 1,2-carboxylate shift. During the first process (**1** $\leftrightarrow$ **2**), the terminal carboxylate of Fe¹ is monodentately coordinated with the metal center, while during the second process (**3** $\leftrightarrow$ **4**) this carboxylate coordinates bidentately (chelating) with the Fe¹-center. In other words, during the first process (**1** $\leftrightarrow$ **2**) the coordination numbers of the Fe centers change from (5C, 5C) in **1** to (5C, 4C) in **2**, while during the second process (**3** $\leftrightarrow$ **4**) the coordination numbers of the Fe centers change from (6C, 5C) in **3** to (6C, 4C) in **4**. Meantime, vertically in Figure 3, the processes **1** $\leftrightarrow$ **3** and **2** $\leftrightarrow$ **4** correspond to a monodentate (top) $\leftrightarrow$ bidentate (bottom) rearrangement of the terminal carboxylate ligand of the Fe¹-center. During these processes, the coordination number of Fe¹ increases by one, from 5C in **1** (and **2**) to 6C in **3** (and **4**).

First, we discuss processes **1** $\leftrightarrow$ **2** and **3** $\leftrightarrow$ **4**, corresponding to the 1,2-carboxylate shift. Comparison of the energies of the **1** and **2** pairs and of the **3** and **4** pairs

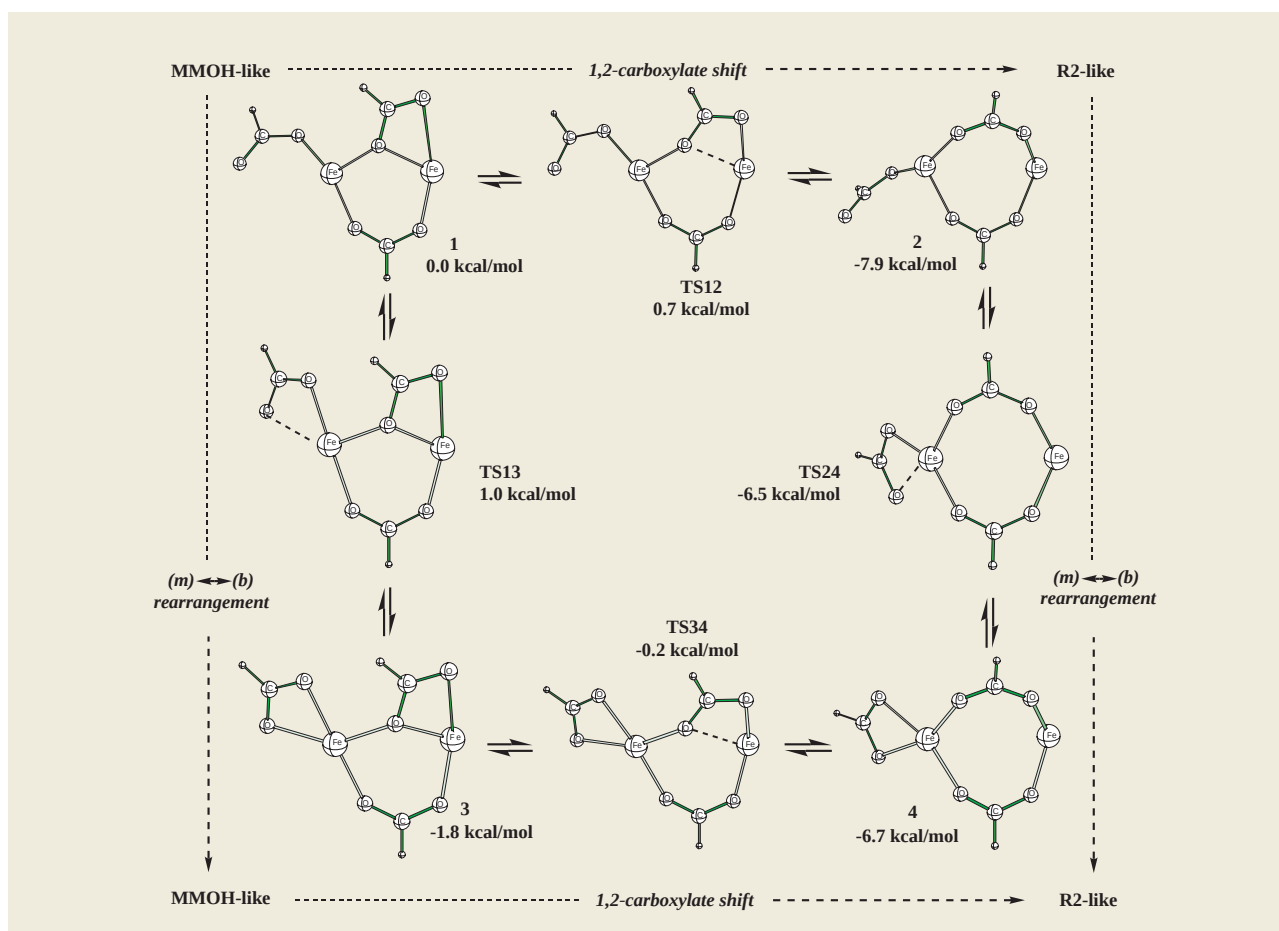


Figure 3

Schematic representation of the reaction mechanism and the energetics of 1,2-carboxylate shift (horizontal) and monodentate $\leftrightarrow$ bidentate rearrangement (vertical) in MMOH-like and R2-like systems. For clarity, ligands not directly involved have been omitted.

(Figure 3) shows that R2-like structures **2** and **4**, with lower coordination numbers, are energetically preferred over their MMOH-like analogs, **1** and **3**; structure **2** is calculated to be 7.9 kcal/mol more stable than structure **1**. Similarly, structure **4** is 4.9 kcal/mol more stable than structure **3**. The obtained difference in the energies of processes $1 \leftrightarrow 2$ and $3 \leftrightarrow 4$ can also be due to the ligand environment of the Fe centers; indeed, in the first process we go from (5C, 5C) in **1** to (5C, 4C) in **2**, whereas in the second process we go from (6C, 5C) in **3** to (6C, 4C) in **4**. Starting from the most stable (**2**) with the lowest (5C, 4C) coordination, one extra coordination on Fe^1 gives 1.2 kcal/mol of destabilization for **4** with (6C, 4C), one extra coordination on Fe^2 gives 7.9 kcal/mol of destabilization for **1** with (5C, 5C), and two simultaneous extra coordinations give 6.1 kcal/mol for **3** with (6C, 5C), which is lower than the sum of the two single coordinations (1.2 + 7.9) by 3.0 kcal/mol due to a cooperative effect.

These changes in the ligand environment also have an effect on the metal–metal distance. From MMOH-like structures **1** and **3** (left) to R2-like structures **2** and **4** (right), the Fe^1 – Fe^2 distance increases by $\sim 0.5 \text{ \AA}$ because of the loss of one (carboxylate) bridge. This is in excellent agreement with the difference in the Fe–Fe distance for the crystallographically determined structures of MMOH_{red} and R2_{red} [18(c), 20] and also with recent mutagenic studies involving μ -1,1 $\leftrightarrow$ μ -1,2 carboxylate shifts [16, 17].

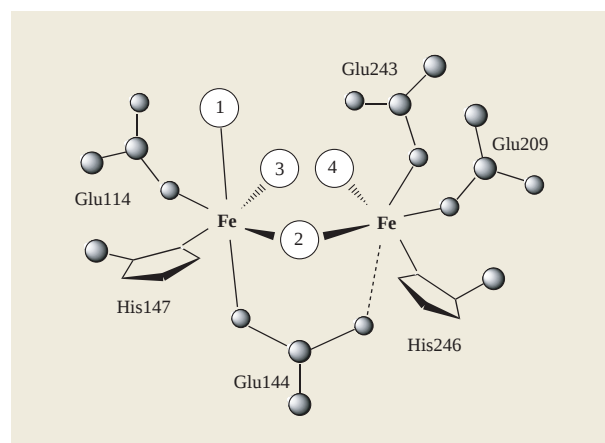
Structures **1** and **2**, as well as **3** and **4**, are separated from each other by small energy barriers, 0.7 and 1.6 kcal/mol, respectively. The transition states corresponding to these barrier heights, **TS12** and **TS34** (Figure 2), each have only one imaginary frequency, 88i and 84i cm^{-1} , respectively, which, according to our normal mode analysis, correspond primarily to the 1,2-carboxylate shift. The first process ($1 \rightarrow 2$), during which Fe^1 remains 5C while Fe^2 changes its ligand environment from 5C to 4C,

is kinetically and thermodynamically more favorable than the second process ($3 \rightarrow 4$), during which Fe^1 remains 6C but Fe^2 changes its ligand environment again from 5C to 4C. Our calculations clearly show that rearrangements $1 \rightarrow 2$ and $3 \rightarrow 4$ are exothermic processes and could take place with very small energy barriers. The key point of our calculations is that the 1,2-carboxylate shift involved in MMOH_{red} and R2_{red} structures is an easy process both thermodynamically and kinetically. This is consistent with the experimental observations described above [18–22].

Next, let us discuss the monodentate $\leftrightarrow$ bidentate rearrangement of the terminal carboxylate ligand in Fe^1 , i.e., processes $1 \leftrightarrow 3$ and $2 \leftrightarrow 4$ for MMOH -like (left) and R2 -like (right), respectively. As one can see from Figure 3, for an MMOH -like structure, the carboxylate ligand tends to be chelating (structure 3) rather than monodentate terminal (structure 1); the process $1 \rightarrow 3$ is calculated to be slightly exothermic by 1.8 kcal/mol, and occurs with a small (1.0 kcal/mol) energy barrier at transition state **TS13**, calculated from **1**. We expect that small variations in the reaction conditions (pH, temperature) may easily reverse the preference observed so far. The computed energy difference (1 kcal/mol) is rather small, and the important fact is that the interconversion is nearly thermodynamically neutral and easily takes place kinetically because of the small barrier.

On the contrary, in the case of R2 -like structures, the carboxylate ligand modeling Asp84 tends to be monodentate terminal (**2**) rather than chelating (**4**) to the Fe^1 -center. The process $2 \rightarrow 4$ is found to be slightly endothermic by 1.2 kcal/mol, and takes place with a 1.4-kcal/mol barrier at the transition state **TS24** (Figure 3). The present result agrees well with experimental observations [6] showing that in the reduced form of R2 , the terminal ligand Asp84 tends to be monodentately coordinated with the Fe^1 -center.

From these results and discussions, we may conclude that the MMOH -like (**1** and **3**) and R2 -like (**2** and **4**) structures are energetically very close to each other and are separated by only small energy barriers [29]. Structures **2** and **4** (corresponding to R2_{red}) with lower coordination numbers in the Fe^2 -center are found to be a few kcal/mol more stable than those with higher coordination numbers (**1** and **3**, respectively). Thus, both the 1,2-carboxylate shift (i.e., the shift of one of the bridging carboxylate ligands from μ -1,1 to μ -1,2 bridging mode between the two Fe centers) and the monodentate $\leftrightarrow$ bidentate rearrangement of terminal carboxylate ligands take place very easily in these systems. Moreover, these reactions can take place *reversibly* under proper experimental conditions. Our results are consistent with the available experimental data [18–22] showing high flexibility of the carboxylate ligands, as well as a nonrigid core structure, in MMOH_{red} and R2_{red} .



Scheme 4

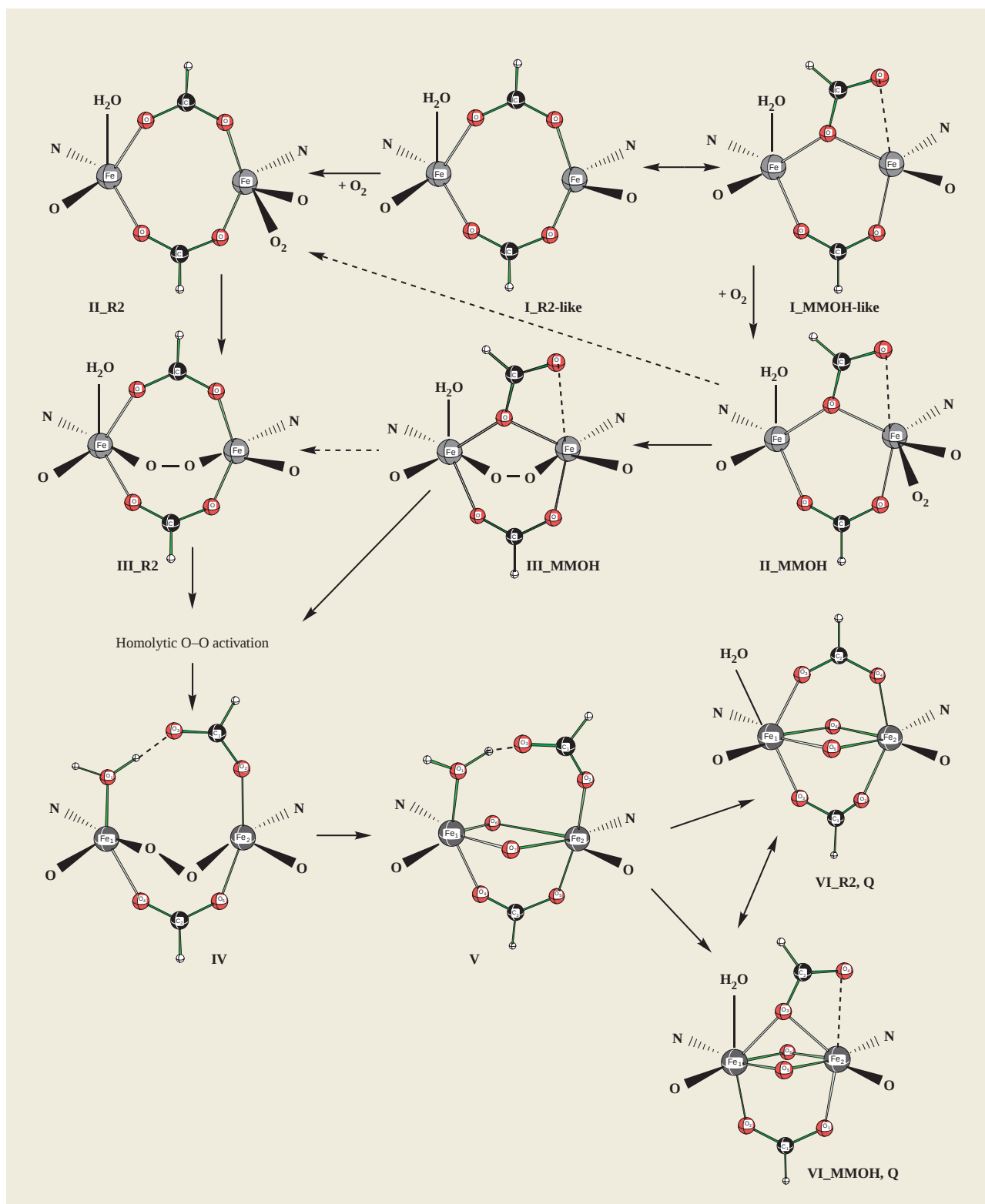
Experimentally proposed possible substrate coordination sites.

Mechanism of dioxygen coordination and O–O bond activation on MMOH_{red} and R2_{red}

Dioxygen coordination and O–O bond activation on MMOH_{red} and R2_{red} are important steps in the conversion of methane to methanol and the formation of a stable Tyr radical, respectively. This is a nontrivial, multistep process. X-ray studies have revealed [1] that the dioxygen molecule most likely attacks the enzyme from the O-ligand side (opposite the protein residues His147 and His246) because of the existence of the substrate coordination pocket. In the literature, four different coordination positions for the O_2 molecule were discussed (see **Scheme 4**) [18, 36, 37]. As discussed above, MMOH_{red} reacts very rapidly with O_2 and forms compound **O**, which is spontaneously converted to compound **P**. Compound **P** is spontaneously converted to compound **Q** with two antiferromagnetically coupled Fe^{IV} centers. While spectroscopic studies [24] indicate that compound **P** is a peroxide species in which both oxygen atoms are bound symmetrically to the iron atoms, the specific coordination mode of O_2 still remains disputable.

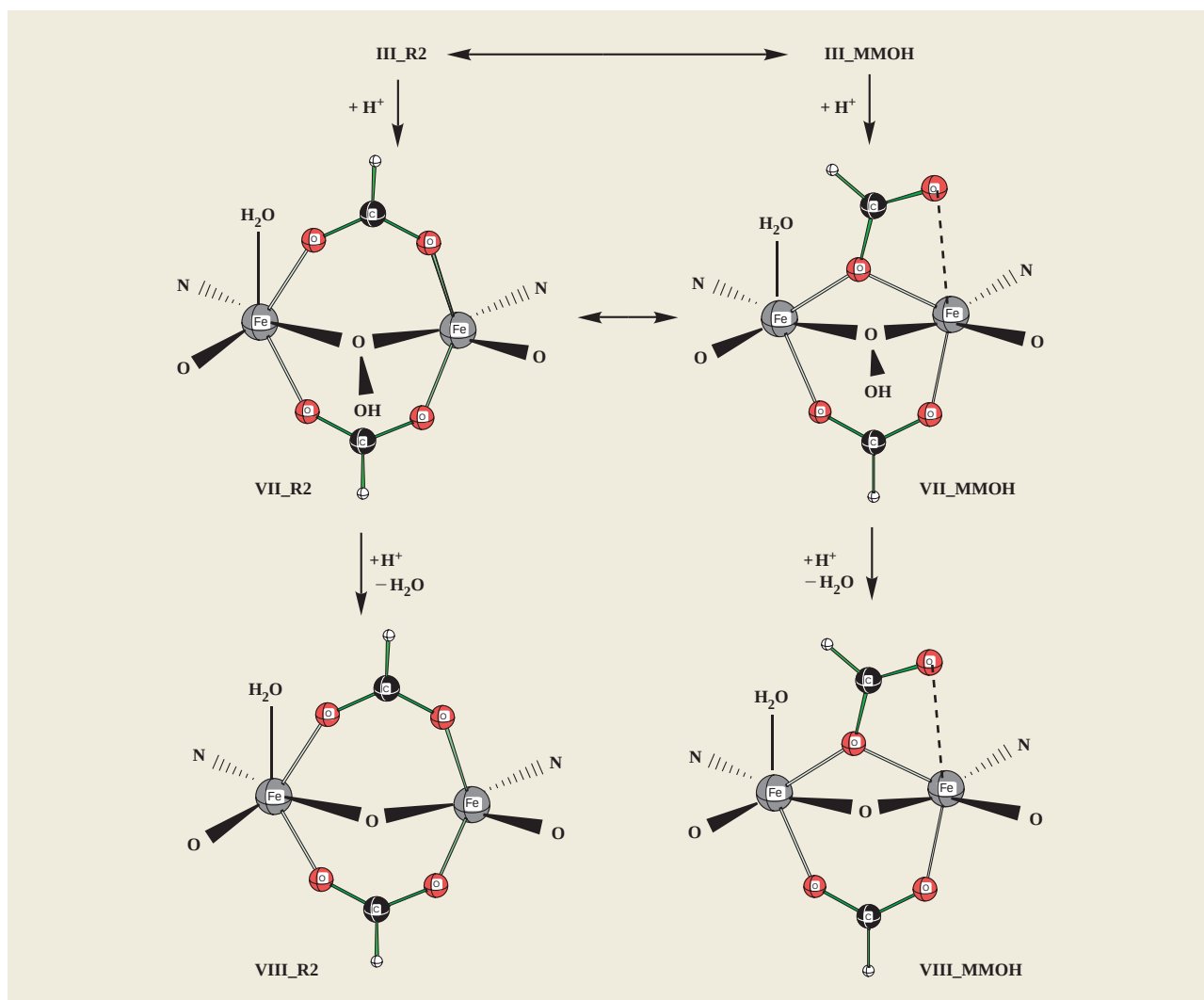
On the basis of available experimental data, as well as our results presented above, which indicate that a) MMOH -like and R2 -like complexes differ ONLY in the coordination mode of one of two bridging carboxylate ligands, and b) R2 -like structures are a few kcal/mol more favorable than MMOH -like structures, and those can be converted to one another with a small energy barrier, we have proposed (see **Schemes 5** and **6**) the mechanisms for dioxygen coordination to MMOH -like and R2 -like complexes, as well as homolytic and heterolytic O–O bond activation in these compounds.

As seen in Scheme 5, the dioxygen molecule can coordinate with either MMOH -like or R2 -like complexes



Scheme 5

Proposed mechanism for the dioxygen coordination and homolytic O-O bond cleavage on MMOH_{red} and R2_{red}. Here, N- is the nitrogen-bearing ligand, His, while O- represents oxygen-bearing ligands such as Glu and Asp.



Scheme 6

Proposed mechanism for the *heterolytic* O–O bond cleavage on MMOH_{red} and R2_{red} . Here, N– is the nitrogen-bearing ligand, His, and O– represents oxygen-bearing ligands such as Glu and Asp.

I to give O-type complexes (**II_MMOH** and **II_R2**, respectively). One may expect that these complexes, like their respective starting complexes **I_MMOH** and **I_R2**, are energetically close, separated by a small barrier, and can easily rearrange to one another and into peroxo-type complexes (**III_MMOH** and **III_R2**, respectively), in which the O_2 molecule is symmetrically coordinated with two Fe-centers in the *cis- μ -1,2* manner. Structures **III** are, most likely, the intermediate **P** proposed by experimentalists. From these complexes, the reaction may proceed via the two different mechanisms (homolytic and heterolytic) presented in Schemes 5 and 6, respectively.

The first step in the homolytic O–O activation mechanism is O–O bond cleavage, followed by the

insertion of O-atoms between the two Fe centers to form $\text{Fe}(\mu\text{-O})_2\text{Fe}$ -type bis- μ -oxo-complexes **V**. Our studies indicate that this process occurs with a large activation barrier [38] if during the reaction all Fe–O(bridging carboxylate) bonds remain intact. However, the existing terminal water ligand on the Fe^1 center facilitates the opening of one of the “legs” of the bridging carboxylate ligand on **III_R2**, and leads to the formation of complex **IV**. Similarly, structure **IV** can easily be generated from intermediate **III_MMOH** by transferring the μ -1,1-carboxylate from the $\text{Fe}^1(\mu\text{-O})\text{Fe}^2$ position to the terminal water molecule. Later, intermediate **IV** rearranges to complex **V** with a small O–O bond activation barrier. In its turn, complex **V** rearranges to the Q-type structures

VI_MMOH and VI_R2 by transferring one of the carboxylate legs from the terminal water ligand to a μ^2 -O position between two Fe-centers with a weak interaction with the Fe^I center.

Support for the heterolytic O–O bond cleavage (see Scheme 6) comes primarily from experimental studies [39] which show that two protons are required in the O₂ reaction, one for the formation of the P-type intermediates (III_MMOH and III_R2) and the other for the conversion of the P-type intermediate to the Q-type intermediates. Lipscomb and co-workers [1, 40] have proposed that a water molecule dissociates as the O–O bond of the P-type intermediate is heterolytically cleaved in the formation of the Q-type structures.

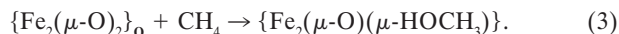
In order to study in detail the mechanisms proposed above, one must formulate the expected spin states of the compounds O, P, and other (not observed) intermediates. The ground state of the O₂ molecule is a triplet ³Σ state with two unpaired electrons. In addition, the ground electronic state of the other reactants, complexes I_MMOH and I_R2, has eight unpaired electrons each from the two ferromagnetically coupled diferrous Fe-centers. Therefore, one may expect that the O-type compounds II_MMOH and II_R2 are a result of the interaction of triplet O₂ with nonet I_MMOH and I_R2. The resultant complex, [Fe^{II}(O₂)Fe^{II}], may have various electronic states, among which the ⁷A, ⁹A, and ¹¹A states should be energetically most favorable.

As expected, the interaction of the triplet O₂ molecule with nonet I_MMOH and I_R2 is extremely weak, and most likely leads to a very shallow minimum corresponding to an ¹¹A state of [Fe^{II}(O₂)Fe^{II}]. The latter could convert almost spontaneously to either mixed-valence species [Fe^{II}(O₂[−])Fe^{III}] or the compound [Fe^{III}(O₂^{2−})Fe^{III}], respectively, via extremely fast one- or two-electron transfer from the Fe-centers (d_π-orbitals) of II_MMOH and II_R2 to the O₂ molecule (π* orbitals). Several electronic states, such as ⁷A, ⁹A, and ¹¹A, can be formulated for the mixed-valence species [Fe^{II}(O₂[−])Fe^{III}], among which ¹¹A is found to be the ground state, while the ⁹A state lies very close to the ¹¹A state. As discussed above, in intermediate [Fe^{III}(O₂^{2−})Fe^{III}], which is most likely the compound P, the O₂ molecule is coordinated simultaneously with two metal centers. The resultant species is [Fe^{III}(O₂^{2−})Fe^{III}].

Our results [29] indicate that ¹¹A is the ground state; ⁹A and ⁷A lie, respectively, 6.7 and 12.0 kcal/mol above ¹¹A. Spin-density analysis shows that intermediate III in its ¹¹A state is clearly an Fe^{III}–Fe^{III} species, with approximately five spins on each Fe-center (4.87 Fe^I and 4.89 Fe²), and some spin delocalized into the two oxygen atoms (0.19 O¹ and 0.05 O²).

Mechanism of the reaction of compound Q with methane

In this section we discuss the reaction



For this purpose we chose the smallest model of compound Q, *cis*-(H₂O)(NH₂)Fe(μ-O)₂(η²-HCOO)₂Fe(NH₂)(H₂O), IX. Note that while spectroscopic studies show that diferryl compound Q contains two AFSC high-spin Fe^{IV} atoms and is diamagnetic with the exchange coupling *J* constant of >60 cm^{−1}, for simplicity we performed spin-unrestricted open-shell single-determinant calculations for the FSC high-spin states ⁹A and ¹¹A. The geometries of the reactants, intermediates, and transition states are shown in Figure 4.

Reactant complex, IX

As seen from Figure 4, the calculated geometric parameters and the general structural character of compound IX, which can be formally written as L₄Fe(μ-O)₂FeL₄, are consistent with experimental findings [1, 4]. Namely, it has two “μ-oxo” (or so-called “diamond”-core oxygen) atoms and two bidentate carboxylate ligands coordinated with the Fe centers. Furthermore, the diamond-core Fe₂O₂ of structures IX, as in the experimentally reported compound Q, has an asymmetric structure; one of the diamond-core O atoms is located closer to one Fe center, and the other closer to the other Fe center. The calculated diamond Fe–O distances in ⁹A are a better match to the experimental values than ¹¹A.

Methane-Q complex, X

The next step of the reaction is expected to be coordination of the incoming substrate with the compound IX. As seen in Figure 4, in general, two distinct pathways, O-side and N-side, corresponding to the substrate coordination with the bridging oxygen atoms, O¹ and O², located on the H₂O- and NH₂-sides, respectively, are possible. Similarly, according to experimental data [1], the only valid pathway is coordination of the substrate from the O-side, because of the existence of the substrate coordination pocket; coordination of the substrate from the N-side is sterically hindered and is not available. In spite of that, we investigated only the N-side pathway. We are currently studying the reaction pathway on the O-side, which will be reported separately. Here, let us suggest that the O-side pathway appears qualitatively similar to the N-side pathway.

The first step of the reaction is expected to be coordination of the incoming substrate with compound IX. Our extensive search shows that coordination of CH₄ with the diamond O-atom is more favorable than any other potentially possible coordination mode. As seen in Figure 4,

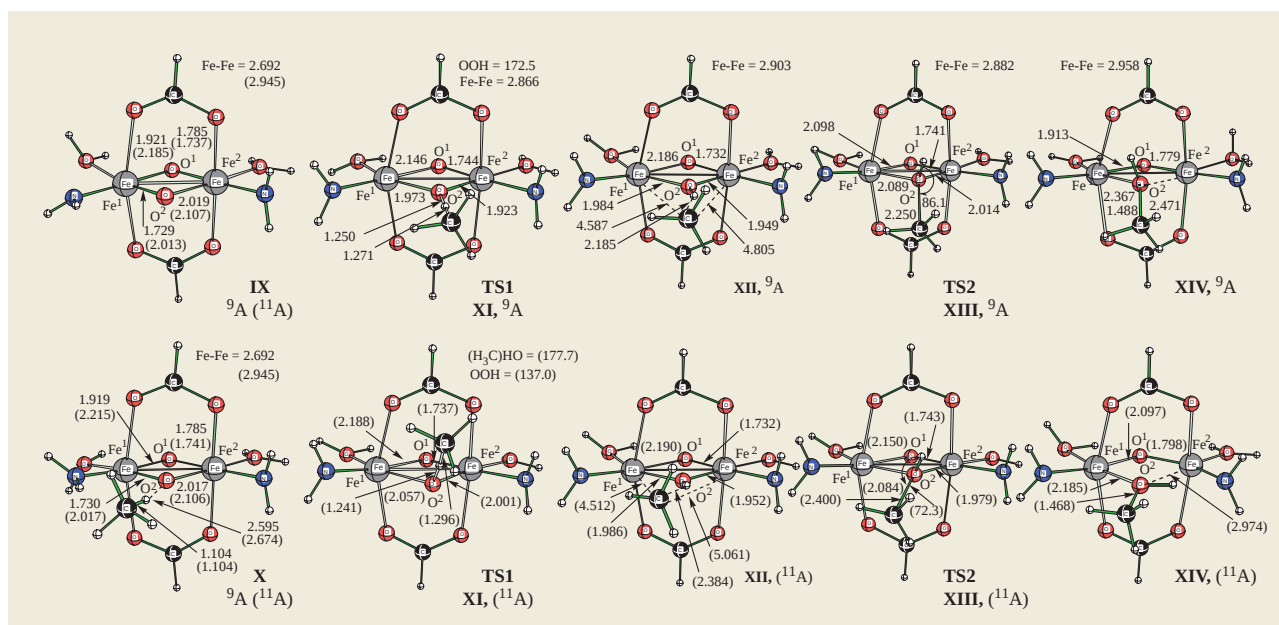


Figure 4

Calculated structures (distances in Å, angles in degrees) of the intermediate and transition states of the reaction: $(\text{NH}_2)(\text{H}_2\text{O})\text{Fe}(\mu\text{-O})_2(\eta^2\text{-HCOO})_2\text{Fe}(\text{NH}_2)(\text{H}_2\text{O}) + \text{CH}_4 \rightarrow (\text{NH}_2)(\text{H}_2\text{O})\text{Fe}(\mu\text{-O})(\mu\text{-HOCH}_3)(\eta^2\text{-HCOO})_2\text{Fe}(\text{NH}_2)(\text{H}_2\text{O})$. Numbers for the ^9A state are without parentheses and those for the ^{11}A state are in parentheses.

coordination of the CH_4 with compound **IX** leads to the methane-Q complex, structure **X**. Since the interaction between CH_4 and structure **IX** is extremely weak (0.7 and 0.3 kcal/mol for the ^9A and ^{11}A states, respectively), it is probable that complex **X** does not exist in reality.

The activation of the methane C–H bond is found to take place on the diamond oxygen O^2 . Our extensive attempts to find other possible pathways for methane C–H activation either converged to the transition states **XI** or led to the transition states higher in energy than structures **XI**. As seen from Figure 4, in transition state **XI** the C–H bond to be broken is elongated from 1.104 Å in **X** to 1.271 Å and 1.296 Å for the ^9A and ^{11}A states, respectively. Furthermore, the O–H bond is nearly formed, with 1.250 Å and 1.241 Å at the transition states, compared to 0.983 Å and 0.978 Å in the products **XII** for the ^9A and ^{11}A states, respectively. These geometrical parameters clearly indicate that **XI** is the transition state corresponding to the H-abstraction process. The H-abstraction barriers are calculated to be 19.5 and 14.4 kcal/mol relative to the corresponding CH_4 complex **X** for the ^9A and ^{11}A states, respectively. These values of the barrier are in reasonable agreement with available experimental estimates, 14–18 kcal/mol [27].

The spin densities in Table 1 for compounds **XI** and **XII** are nearly identical for ^9A and ^{11}A , except for those on

the $\text{O}^2\cdots\text{H}\cdots\text{CH}_3$ fragment. For the CH_3 group, the spin densities for ^9A and ^{11}A are -0.46 and $+0.52$, respectively.

Overcoming the transition state **XI** leads to the product **XII**, in which the Fe centers are bridged by one O and one OH ligand. **XII** is an oxo-hydroxyl complex, formally written as $\text{L}_4\text{Fe}(\mu\text{-O})(\mu\text{-OH})\text{FeL}_4$, with the methyl radical only weakly interacting via a $\text{C}\cdots\text{HO}$ interaction.

As seen in Table 1, in the intermediate **XII**, the CH_3 group is now a bound radical with spin densities of -0.98 and 1.00 for the ^9A and ^{11}A states, respectively. The spin densities on Fe^1 and Fe^2 in **XII** can qualitatively be considered to correspond to Fe^{IV} with four spins and Fe^{III} with five spins, for the ^9A and ^{11}A states, respectively.

The spin densities of the complex **IX**, transition states **XI**, and products **XII** presented in Table 1 can be interpreted in the following way. In structure **XI**, the total of eight and ten unpaired electrons (spins) of the ^9A and ^{11}A states, respectively, are equally localized on two L_nFeO groups. According to the discussions presented above, one may consider **IX** (and **X**) in the ^9A state as an $[\text{Fe}^{\text{IV}}\text{-Fe}^{\text{IV}}]$ complex, and **IX** (and **X**) in the ^{11}A state as an $[\text{Fe}^{\text{IV}}\text{-Fe}^{\text{III}}]$ mixed-valence complex. In both transition states **XI**, a radical center begins to develop on the CH_3 group, with spin densities of -0.46 and $+0.52$ for the ^9A and ^{11}A states, respectively. In the intermediate **XII**, the CH_3 group becomes a bound radical with spin densities of

Table 1 Mulliken atomic spin densities for the various intermediates and transition states of the MMO + methane reaction, calculated at the B3LYP/SBK level.

Structures	Atomic spin densities (e)					
	Fe^1	Fe^2	O^1	O^2	H^*	CH_3^{**}
<i>⁹A states</i>						
IX	3.17	3.37	0.33	0.45	—	—
X	3.16	3.36	0.31	0.45	—	—
XI	3.40	4.09	0.56	−0.22	0.03	−0.46
XII	3.32	4.15	0.59	0.19	0.00	−0.98
XIII	3.28	4.14	0.50	0.21	0.00	−0.81
XIV	2.80	4.15	0.35	0.04	0.00	0.00
<i>¹¹A states</i>						
IX	3.35	4.10	0.57	1.21	—	—
X	3.37	4.10	0.58	1.21	—	—
XI	3.35	4.12	0.59	0.74	−0.05	0.52
XII	3.34	4.14	0.59	0.22	0.02	1.00
XIII	3.61	4.14	0.55	0.10	0.00	0.85
XIV	4.10	4.13	0.57	0.05	0.00	0.00

*H atom located between O^2 and CH_3 fragments.

**Number for the entire CH_3 fragment.

−0.98 and 1.00 for the 9A and ^{11}A states, respectively. Now, the spin densities on Fe^1 and Fe^2 in **XII** can qualitatively be considered to correspond to Fe^{IV} and Fe^{III} for the 9A and ^{11}A states, respectively. In going from **X**(9A) to **XI**(9A), the formal oxidation state of Fe^2 changes from Fe^{IV} to Fe^{III} ; going from **X**(^{11}A) (which is already Fe^{III}) to **XI**(^{11}A), no such change is required. Since the two Fe centers are coupled ferromagnetically in both the 9A and ^{11}A states, the spin of the CH_3 radical in both **XI** and **XII** must couple antiferromagnetically (with negative spin) and ferromagnetically (with positive spin) to make the total spin, $2S + 1$, equal to 9 and 11, respectively. In the bound radical complexes **XII**(9A) and **XII**(^{11}A), in which the interaction of the CH_3 radical with the two iron atoms is very weak, the total energies are nearly identical, and the optimized geometries in **Figure 5** are also nearly identical. In the transition states **XI**(9A) and **XI**(^{11}A), in which the spin is about 50% developed, the energies and geometries are similar. The fact that the system tends to adopt the same structures, whether the overall spin state $2S + 1 = 9$ or 11, suggests that this mixed-spin [$Fe^{IV}-Fe^{III}$] electronic configuration is the preferred state of intermediate **XII**. The present spin-density analysis clearly demonstrates that the methane oxidation proceeds via a bound-radical mechanism, which is in good qualitative agreement with experiments of Valentine et al. [41].

The entire process $IX + CH_4 \rightarrow XII$ is calculated to be endothermic by 11.4 and 3.0 kcal/mol for the 9A and ^{11}A states, respectively.

The next step in the reaction path should be the reaction of the methyl radical with the complex, which is found to occur via TS **XIII**. Figure 4 reveals that this transition state involves a torsion motion of the hydroxyl OH ligand before the methyl radical can add to the bridging hydroxyl ligand to form methanol. As the methyl radical approaches the hydroxyl ligand O^2H from below the core $Fe^1O^1Fe^2O^2$ plane, the H atom of the hydroxyl ligand leaves the coplanar tricoordinate O environment and bends upward to create a tetrahedral tetracoordinate O environment. Apparently this bend costs some energy before it is recovered by reaction of the methyl group to form a C–O covalent bond. The barrier heights for the CH_3 addition to the hydroxyl ligand calculated relative to the intermediate **XII** are 6.8 and 5.8 kcal/mol for the 9A and ^{11}A states, respectively. Obviously, this step of the reaction is not rate-determining, and can occur rather rapidly; the lifetime of the radical intermediate is expected to be too short to be detected experimentally.

Overcoming the barrier at **XIII** leads to the methanol complexes **XIV**. The overall reaction $IX + CH_4 \rightarrow XIV$ is calculated to be exothermic by 31.4 and 48.6 kcal/mol for the 9A and ^{11}A states, respectively.

Table 1 shows that in the ^{11}A state, upon going from the **XII** to the **XIV**, Fe^1 changes its formal oxidation state from Fe^{IV} with four spins to Fe^{III} with five spins, while the spin density on the methyl radical is completely annihilated upon forming a covalent bond between CH_3 and OH. The transition state **XIII** has a spin distribution between that of **XII** and **XIV**. On the other hand, in the 9A state, upon going from the hydroxyl complex **XII** to the

methanol complex **XIV**, the spin density on Fe^I is reduced by about 0.5, corresponding to the disappearance of roughly one unpaired electron. Since Fe^V is not a stable species, it is probable that Fe^I changes its formal oxidation state from Fe^{IV} with four spins to Fe^{III} with five formal d electrons. Because of the $2S + 1 = 9$ restriction (i.e., the total number of unpaired electrons must be 8 within the Fe^{III}–Fe^{III} core), Fe^I in **XIV** chose to form one d lone pair with only three spins remaining. This complex **XIV**(⁹A) is thus higher in energy than the corresponding complex **XIV**(¹¹A), in apparent violation of Hund's rule. The final step of the reaction is an elimination of the methanol molecule and regeneration of the enzyme. This step of the reaction could be a complex process; it is currently under study.

Finally, let us discuss the entire potential energy profiles of the reaction: **IX** + CH₄ → **XIV** for both ⁹A and ¹¹A states, as shown in Figure 5. First, the potential energy profile does not differ much between the ⁹A and ¹¹A states. However, there are some differences in the energetics between the two states. At reactant **IX**, the ⁹A state is lower than the ¹¹A state by 8.8 kcal/mol, while at the first transition state **XI**, the energy gap between the ⁹A and ¹¹A states is reduced to 3.7 kcal/mol. The ⁹A state of **IX** has, from a qualitative point of view, a Fe^{IV}–Fe^{IV} core, as suggested experimentally for **Q**, while the ¹¹A state of **IX** has a Fe^{IV}–Fe^{III} core and is less stable. Once the system reaches the hydroxyl complex **XII**, calculations for both ⁹A and ¹¹A converge to the same electronic state with the same structure and energy, corresponding to the Fe^{IV}–Fe^{III} mixed-valence state, which interacts very weakly with the methyl radical either ferro- or antiferromagnetically. In the product methanol complex **XIV**, the ¹¹A state is 8.4 kcal/mol lower than the ⁹A state. Here, the preferred iron core is Fe^{III}–Fe^{III}, and each Fe has five spins, which naturally gives the ¹¹A state when ferromagnetically coupled.

Thus, the results presented above show that the MMO-catalyzed methane hydroxylation reaction proceeds via a *bound-radical mechanism through the rate-determining H abstraction transition state*. During the process, the oxidation state of the Fe core changes from Fe^{IV}–Fe^{IV} in **IX** to a mixed-valence Fe^{IV}–Fe^{III} in the short-lived intermediate **XIV**, and finally to Fe^{III}–Fe^{III} in **XIV:IX**, $L_4Fe^{IV}(\mu-O)_2Fe^{IV}L_4 + CH_4 \rightarrow XI$, TS1 for CH activation (rate-determining) → **XII**, $L_4Fe^{IV}(\mu-O)(\mu-OH(\cdot\cdot CH_3))Fe^{III}L_4$ (bound-radical intermediate) → **XIII**, TS2 for CH₃ addition → **XIV**, $L_4Fe^{III}(OHCH_3)(\mu-O)Fe^{III}L_4$.

Perspective

The above calculations using small and medium-sized models have already provided new insights into the mechanisms of reactions and transformations in the MMO

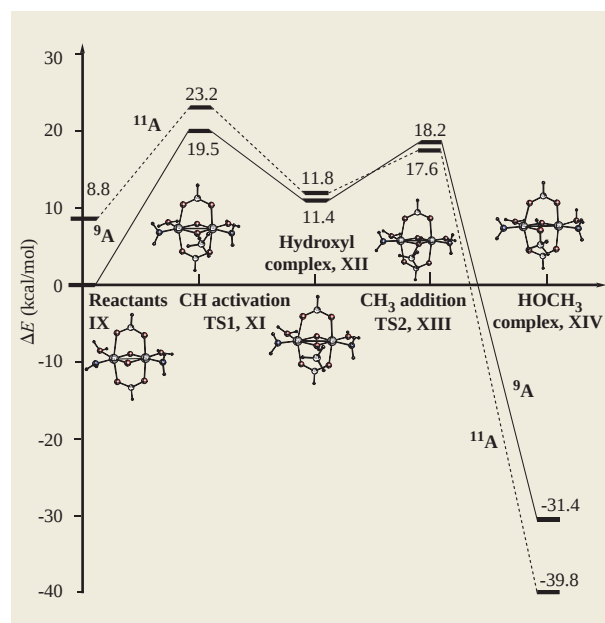


Figure 5

Potential energy profile (in kcal/mol) at the B3LYP/SBK(CO*)/B3LYP/SBK level for both the ⁹A state and the ¹¹A state of the methane activation reaction: $(NH_2)(H_2O)Fe(\mu-O)_2(\eta^2-HCOO)_2Fe(NH_2)(H_2O) + CH_4 \rightarrow (NH_2)(H_2O)Fe(\mu-O)(\mu-HOCH_3)(\eta^2-HCOO)_2Fe(NH_2)(H_2O)$.

and R2 systems. This is because these processes are essentially electronic in nature, and the understanding of the behavior of electrons (how the spin state changes, how the substrate structure changes, and how the potential energy varies) during the processes is essential for understanding of the mechanism. However, we recognize that these simple models do not describe real metalloenzyme systems, including their entire protein environment. For instance, the existence of different amino acids on the active sites of MMOH and RNR has not been considered. To study the effects of protein environment on the catalytic activities of enzymatic systems, one must consider all (or many) of the ligands and water molecules in the real enzymes. There exists virtually no modeling calculation of enzymatic reactions in which the effects of protein environment are fully taken into account. How we accomplish this is the most serious problem to be solved, and we return to it later in Section 4.

3. The ONIOM method—A new computational tool for the study of large molecular systems

The electronic structure theory used in the above calculations for models of metalloenzyme systems is a density functional theory (DFT) method with the B3LYP functional. There exists a hierarchy of theoretical methods

Table 2 Theoretical methods for calculating molecular structure and the energy of molecular systems.

<i>Approximation</i>	<i>Reliability</i>	<i>Size dependency</i>	<i>No. of non-H atoms</i>
<i>Ab initio MO methods</i>			
CCSD(T)/QTZ	Quantitative (~2 kcal/mol)	$\sim N^6$	5–6
MP2/DZP	Semiquantitative	$\sim N^4$	10–50
DFT(B3LYP)/DZP	Semiquantitative	$\sim N^3$	50–100
HF/DZ	Qualitative	$\sim N^3$	100–200
<i>Semiempirical MO methods</i>			
AM1, PM3, MNDO	Semiquantitative	$\sim N^{2-3}$	1000
<i>Classical force-field methods</i>			
Amber, Charmm, MM3	Semiquantitative	$\sim N^{1-2}$	10^{3-4}

for calculations of the molecular structure and energy of molecular systems, as shown in **Table 2**.

For instance, the CCSD(T) method (coupled cluster with single and double excitations with perturbation correction for triples), one of the most reliable *ab initio* molecular orbital (MO) methods, can calculate the energetics, such as the enthalpy of formation, with an average error from the experiment of 2 kcal/mol or so, and provides what often is called “chemical accuracy,” the quantitative accuracy needed for chemistry. However, this method is very expensive, and the largest molecule for which this method can be routinely applied is limited at present to one containing five to seven nonhydrogen atoms. Its computer time requirement increases approximately in the order of $\sim N^6$, where N is the number of atoms; with this high- N dependence, the maximum size of an applicable molecule increases very slowly with the increased speed of available computers. This quantitative method cannot be applied directly to enzyme reactions in the foreseeable future. In addition, some semiquantitative methods, such as MP2/DZP or DFT(B3LYP)/DZP, give semiquantitative results with which the mechanisms of chemical reactions can be elucidated quite well, although the calculation of energies at this level of approximation is not in general good enough for determination of the absolute value of the rate constant. Some semiempirical MO methods, such as AM1 and PM3, are fast but provide semiquantitative results (i.e., conclusions derived from these methods are usually acceptable, with some exceptions). At the end of the hierarchy are the empirical classical molecular mechanics (MM) or force-field methods, such as MM3, Charmm, and Amber. These MM methods in general express molecular energy as a sum of simple functions of bond distances, bond angles, and torsion angles. With parameters fitted to experimental values, calculations are extremely rapid and give reasonable molecular and intermolecular structures

and energies for systems for which the parameters are calibrated. One of the most serious drawbacks of the MM methods is that they usually cannot describe the breaking or formation of bonds (i.e., chemical reactions). Therefore, if one is interested in chemical reactions, one must use DFT (with good density functional) or MP2 level for qualitative results, and better methods for more quantitative results.

In recent years, computational chemists have become interested in chemical reactions of large molecular systems. As discussed in Section 3, for instance, we are interested in enzymatic reactions, chemical reactions in the protein environment. In other cases, one is interested in a series of organometallic reactions in homogeneous catalysis, in which the catalyst is typically a transition-metal complex with complicated ligands. In yet other cases, one is interested in chemical reactions in solution, in which some solvent molecules participate directly in the reaction or form site-specific complexes, while many other solvent molecules also affect the reaction mechanism and rate.

Using present and foreseeable future technology, it is not likely that a reliable high-level MO method can be applied to the entire molecular systems considered above. Theoreticians usually have two options. One is to adopt a simple model system to which one can apply a reasonable theoretical method. In Section 2, we did exactly this, by adopting models consisting only of metal atoms, some simplified ligands, and substrate. However, by using simplified models, one misses the effects of the protein environment, the effects of real complex ligands, or the effects of solvent molecules. These effects might change the transition-state structure and energetics, and consequently the mechanism of reaction. The second option is to adopt an inexpensive and not very reliable method with which to perform the calculation for the real (or more realistic) system. If cheap methods are used, one's results and conclusion may not be correct.

However, one can easily recognize that the chemical reaction and many other chemical problems are often a local phenomenon. A chemical reaction typically involves a few to several atoms in the action center (the core of an onion), with bond formation and breaking taking place among them. Obviously, one must use an expensive, reliable method to describe this active region of the system. The atoms in the vicinity of the action center (the second layer of the onion) participate in the reaction indirectly by pulling or pushing electrons; this region probably has to be handled with an appropriate MO method in order to take into account its electronic effects. Atoms farther away from the action center (the third layer or outer skin of the onion) are likely to provide some field, such as the long-range electrostatic interaction or nonbonding steric interaction; they can probably be treated with an inexpensive MO method or a molecular mechanics method. This kind of consideration leads to the idea of a hybrid method, a theoretical method that uses different methods for different parts of the molecular system.

The hybrid method has in fact been in use for some time. In particular, the QM/MM methods, which combine an MO method and an MM method, have been used successfully for the modeling of biological systems. These methods are usually combination-specific; i.e., a specific MO method (primarily semiempirical) and a specific MM method are combined, and additional parameters are often introduced for better agreement with experiment. We refer the readers to a recent review paper for a discussion of these methods [42].

We have recently developed the ONIOM method. At first, we used the *integrated molecular orbital + molecular mechanics* method (IMOMM), which integrates MO methods with the MM3 force field [43] and has a methodology slightly different from QM/MM. Soon afterward, it was realized that the extrapolation scheme in IMOMM could be generalized to combine two MO methods as well. This resulted in a combined MO + MO method, which was referred to as the *integrated molecular orbital + molecular orbital* method (IMOMO) [44]. Later, the integration of more than two methods was accomplished, and the entire suite of integrated methods was named the ONIOM method (*our own n-layered integrated molecular orbital + molecular mechanics method*) [45]. Thus, IMOMO encompasses both two-layered ONIOM2(MO:MO) and three-layered ONIOM3(MO:MO:MM), and IMOMM is in principle equivalent to ONIOM2(MO:MM) and ONIOM3(MO:MO:MM). The latest incarnation of ONIOM is available in Gaussian98, and can integrate up to three layers of most electronic structure methods and/or molecular mechanics methods from the package [46]. Various aspects of the ONIOM

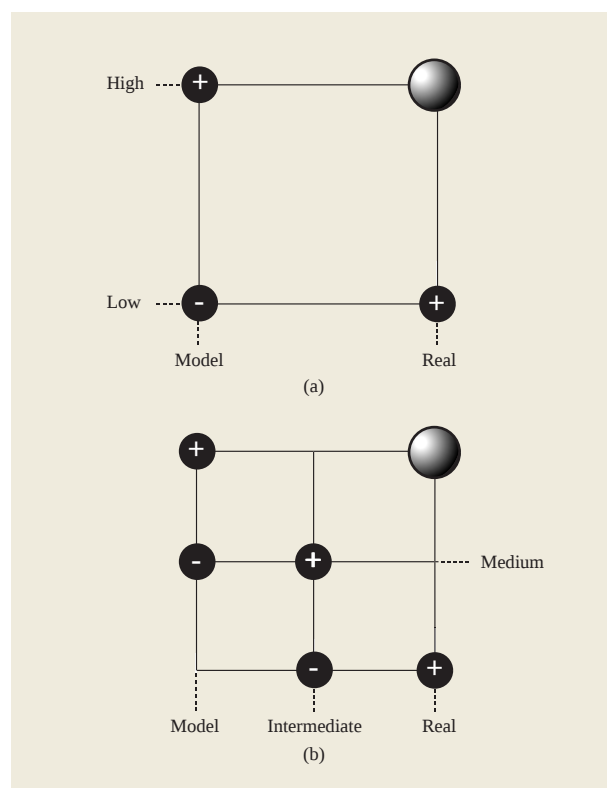


Figure 6

Schematic representation of (a) two-layer and (b) three-layer ONIOM extrapolation schemes.

method have been reviewed previously [47–49]. In the present review, we briefly discuss theory and some applications of the ONIOM method.

Foundation of the ONIOM method

In the two-layered ONIOM method, the total energy of the system is obtained from three independent calculations:

$$E^{\text{ONIOM2}} = E_{\text{model}}^{\text{high}} + E_{\text{real}}^{\text{low}} - E_{\text{model}}^{\text{low}}, \quad (4)$$

where *real* denotes the full system, which is treated at the *low* level, while *model* denotes the part of the system for which the energy is calculated at both *high* and *low* levels. The concept of the ONIOM method is represented schematically in **Figure 6**. One can see that the method can be regarded as an extrapolation scheme. Beginning at $E_{\text{model}}^{\text{low}}$, the extrapolation to the high-level calculation ($E_{\text{model}}^{\text{high}} - E_{\text{model}}^{\text{low}}$) and the extrapolation to the real system ($E_{\text{real}}^{\text{low}} - E_{\text{model}}^{\text{low}}$) are assumed to produce an estimate for $E_{\text{real}}^{\text{high}}$. For a three-layer ONIOM scheme, the energy expression can be written as

$$E^{\text{ONIOM3}} = E_{\text{small model}}^{\text{high}} + E_{\text{intermediate model}}^{\text{medium}} - E_{\text{small model}}^{\text{medium}} + E_{\text{real}}^{\text{low}} - E_{\text{intermediate model}}^{\text{low}} \quad (5)$$

The definition of the model system is rather straightforward when there is no covalent bond between the layers. If one considers a solute molecule solvated with one solvent molecule to be the *real* system, the solute will be the *model* system, and the solute–solvent interaction is naturally included in the *low*-level calculation of the *real* system. The situation becomes more complicated when a covalent bond exists between the layers. The most straightforward way to ensure that the model system is representative of the real system is to saturate the dangling bonds with *link atoms*, which is the scheme chosen for the ONIOM method. For example, if a methyl group is treated at the low level, a hydrogen atom can be substituted for this group in the model system. One chooses link atoms that best mimic the substituents that exist only in the real system. In practice, hydrogen atoms are good link atoms when carbon–carbon bonds are broken, but in some cases other atoms may be better [50]. One may also shift the one-electron energy of the link atom with a shift operator to mimic the electronegativity of the replaced atom [51], although we have not explored its usefulness in ONIOM in detail.

As can be seen, there is no restriction on the methods used at various levels (high, medium, low), and various MO and MM combinations discussed above can be derived. Combining different levels of MO methods is a unique feature of the ONIOM method. It turns out that MO:MO integration is rather straightforward, and virtually no special attention is required. On the other hand, the integration in ONIOM of MO and MM methods, combining two methods with very different philosophies, leads to many serious problems, as with all of the QM/MM methods. Although we do not discuss these technical problems here, we revisit this problem briefly in Section 4. The integration of two MO levels with one MM level, ONIOM3(MO:MO:MM), is unique, a feature absent from other QM/MM methods. In our MO–MM combinations, the interaction between the MO and MM regions can be treated at the MM level, i.e., with so-called mechanical embedding [52], or, alternatively, in the QM Hamiltonian, with so-called electronic embedding (in the near-future release of Gaussian software).

In the construction of the ONIOM model system, atoms that belong to the high-level layer have the same coordinates as the corresponding atoms in the real system. Even during geometry optimizations, these coordinates remain identical to one another. When no bond exists between the two layers, the first derivative of the energy with respect to the geometry is easy to obtain:

$$\frac{\partial E^{\text{ONIOM}}}{\partial q} = \frac{\partial E_{\text{model}}^{\text{high}}}{\partial q} + \frac{\partial E_{\text{real}}^{\text{low}}}{\partial q} - \frac{\partial E_{\text{model}}^{\text{low}}}{\partial q} \quad (6)$$

However, the link atoms used in the model system do not exist in the real system, and one of the main issues in this type of hybrid method is their geometrical placement, or how the geometry of the model system is related to that of the real system. We assume that the link atoms are connected to the high-level layer with the same angular and dihedral values as the link atom hosts (LAHs, the atoms replaced by the link atoms in the model system) in the real system. Now steric effects of the substituents are also taken into account in the two model system calculations. In our earlier implementation in the IMOMM method, we used fixed (standard) bond lengths between the link atoms and the high-level layer, as well as fixed bond lengths between the LAH atoms and the high-level layer. Although this scheme works well for geometry optimization, one degree of freedom is lost for each link between the high- and low-level layers, which causes problems, for example, with dynamics or frequency calculations.

In the later implementation including Gaussian98 [50], the angles and dihedrals are treated in the same manner as in the IMOMM scheme, while the bond distances between the high-level layer and the link atoms are obtained by scaling the corresponding distances between the high-level layer and the LAH atoms:

$$\mathbf{R}_{\text{link}} = \mathbf{R}_{\text{high-level atom}} + g(\mathbf{R}_{\text{LAH}} - \mathbf{R}_{\text{high-level atom}}), \quad (7)$$

where $\mathbf{R}_{\text{high-level atom}}$ denotes the atom in the high-level layer to which the link atom is connected. The scaling factor g is chosen so that reasonable bond lengths between the LAH atoms and high-level-layer atoms also yield reasonable bond lengths between the link atoms and high-level-layer atoms. In this version, the correct number of degrees of freedom ensures that the potential energy surface is well defined, and gradients and higher derivatives are available [50]. In this case, one must use the Jacobian $\mathbf{J}$ to convert the coordinate system for the model system to the coordinate system for the real system:

$$\frac{\partial E^{\text{ONIOM}}}{\partial q} = \frac{\partial E_{\text{model}}^{\text{high}}}{\partial q} \cdot \mathbf{J} + \frac{\partial E_{\text{real}}^{\text{low}}}{\partial q} - \frac{\partial E_{\text{model}}^{\text{low}}}{\partial q} \cdot \mathbf{J} \quad (8)$$

The Hessian or higher-order derivatives can be uniquely defined in a similar fashion. Any method for the investigation of potential energy surfaces based on conventional techniques can now be used with the ONIOM method.

Bond dissociation energies

ONIOM methods can be applied to virtually any chemical problem, but one of the first areas that has been covered

Table 3 Bond dissociation energies (kcal/mol) using a minimal model. See references for specific procedures and origin of experimental data.

Reaction ^a	Ref. ^b	Experimental	G2MS(R):RMP2 ^c	G2MS(U):RMP2 ^c
Ph-F	I	125.7 ± 2	124.6	125.2
Ph-CH ₃	I	101.8 ± 2	100.3	101.9
PhSiH ₂ -H	I	88.2	86.3	86.3
PhCH ₂ -SCH ₃	I	61.4 ± 2	61.6	62.4
F ₅ SO-OSF ₅	I	37.2 ± 2	37.3	37.7
PhCH ₂ -H	II	88.0 ± 1	90.1	90.4
MePhCH-H	II	85.4 ± 1.5	87.8	88.2
Me ₂ PhC-H	II	84.4 ± 1.5	87.2	87.6
Ph ₂ CH-H	II	84 ± 2	82.6	82.9
MePh ₂ C-H	II	81 ± 2	82.6	82.9
Ph ₃ C-H	II	n.a.	75.9	
PhCH ₂ -CH ₃	II	75.8 ± 1	79.0	79.7
MePhCH-CH ₃	II	74.6 ± 1.5	78.1	78.9
Me ₂ PhC-CH ₃	II	73.7 ± 1.5	77.2	77.9
Ph ₂ CH-CH ₃	II	72 ± 2	72.7	73.5
MePh ₂ C-CH ₃	II	69 ± 2	72.9	73.6
Ph ₃ C-CH ₃	II	n.a.	64.1	
G2MS(R):RMP2:RHF ^d				
C ₆₀ :S ₀ → T ₁ (π bond)	III	36.1	35.1	
G2MS(R):RMP2:B3LYP ^e				
Ph ₃ C-CPh ₃	IV	n.a.	16.4	

^aThe dissociation bond is indicated by the dash.

^bI from [54]; II from [56]; III from [55]; IV from T. Vreven and K. Morokuma, work in preparation.

^cGeometries and frequencies from B3LYP, 6-31G(d) basis set employed for the MP2 calculations.

^dGeometry from AM1, no ZPE correction, intermediate model is naphthalene.

^e6-31G basis set for B3LYP, geometry obtained at ONIOM2(B3LYP/6-31G:B3LYP/3-21G) level; see Figure 2 for partitioning.

in great detail is the calculation of bond dissociation energies (BDEs). In addition to several papers in which BDEs were used to illustrate the method [50, 53], a number of studies have focused exclusively on this topic. The dissociation energies of sets of larger molecules were calculated by Froese [54, 55] and Vreven¹ [56]. In both studies, a number of high and low levels were investigated, but the geometries and zero-point vibrational energies or temperature corrections were obtained at the nonintegrated B3LYP level. In most cases, a minimal model was employed (the non-H atoms only in the dissociation fragment), and the results show that the G2MS:RMP2 combination yields BDEs very close to the experimental values [G2MS [57] is our own inexpensive G2-like method, G2MS(R/U) = CCSD(T)/6-31G(d) + (R/U)OMP2/6-311 + G(2df,2p) - (R/U)OMP2/6-31G(d), which we have used extensively in our IMOMO studies]. In **Table 3** we show the molecules from these studies that contain more than six non-H atoms. The error with the experimental values is usually not more than a few kcal/mol, and could probably be reduced further by fine-tuning the choice of basis sets. The primary conclusion is that a very high-level correlation method is required for

the reaction center, but a low-level correlation method is still needed for the remainder of the system. Spin-contamination in the low-level method is an important problem for molecules containing aromatic substituents, and unrestricted methods have therefore proven inappropriate. We were able to predict the BDE for Ph₃C-H and Ph₃C-CH₃, for which the experimental values are not available.

The last two entries in Table 3, C₆₀ fullerene and hexaphenylethane (HPE), are systems too large for the G2MS:MP2 calculation, and it was necessary to introduce a third layer treated at a less expensive RHF level. In the case of C₆₀, RHF for the lowest level proved useful, and B3LYP was a good choice for HPE. The HPE example is particularly interesting; this compound is thought to be too unstable to exist [58], but our BDE value suggests that synthesis could be possible.

Diels-Alder reactions

In addition to bond dissociation energies, Diels-Alder reactions have been studied extensively with the ONIOM method. In the paper that first defined the ONIOM method, we discussed the addition of acrolein to substituted 1,3-butadienes [45], where ethylene-butadiene

¹ T. Vreven and K. Morokuma, work in preparation.

Table 4 Experimental and IMOMO(G2MS:MP2) activation barriers (kcal/mol) for Diels–Alder reactions.

	G2MS:MP2	Experimental
Butadiene + butadiene	23.5	23.7 ± 0.2, 24.5
Acrolein + isoprene	17.6	18.7
Maleic anhydride + isoprene	9.2	12.2
Maleic anhydride + 2- <i>tert</i> -butyl-1,3-butadiene	4.6	6.5

was used as the model system. It was expected that the electronic effect of the COH group in acrolein must also be treated with an MO method (MM is incapable of treating the electronic effects), and one was therefore included at an intermediate layer. The alkyl substituents on butadiene were not expected to play an important role, and were treated at a third, low-accuracy level. Both the optimization and energy calculations were performed with a variety of combinations. The ONIOM(B3LYP:HF:MM3) combination (the 6-31G basis set has been used throughout) yields a deviation in the activation barrier of only 1.4 kcal/mol from the B3LYP target for both the reaction of acrolein with isoprene (2-methyl-1,3-butadiene) and the reaction of acrolein with 2-*tert*-butyl-1,3-butadiene. The third reaction studied, ethylene + *s-trans trans,trans*-1,4-di-*tert*-butyl-1,3-butadiene, with MP4 as the target method, yields good results with the ONIOM(MP4:MP2:MM3) combination. The error in activation energy is 2.3 kcal/mol, which is significantly less than the error of 4.4 kcal/mol for MP4:HF:MM3. This is similar to the IMOMO studies of bond dissociation energies, in which we observed that results are often inadequate when the HF level is employed for a layer too close to the reaction center.

The second investigation of Diels–Alder reactions was reported in the paper that introduced G2MS [57]. The IMOMO(G2MS:MP2)//B3LYP + ZPE(B3LYP)

combination was very successful, and two subsequent IMOMO studies of Diels–Alder reactions focused on this combination [59, 60]. Because G2MS calculations are too expensive for systems with more than eight non-H atoms, the method was tested by comparing the results directly to experimental values. In **Table 4** we summarize the IMOMO(G2MS:MP2) results for Diels–Alder reactions for which experimental activation energies are available. The computational data is in fairly good agreement with the experimental values. In **Table 5** we show the branching ratio of several reactions. The results do not always agree quantitatively with the experimental values because the branching ratio is very sensitive to small errors in our calculated energies. However, IMOMO(G2MS:MP2) reproduces qualitatively well the reversal of *syn/anti* reactivity for ethylene and acetylene with **1**, which is caused by the repulsion between the lone pair of oxygens in **1** with the π -orbitals of acetylene in the *syn* transition state.

The S-value test

The most important questions when the ONIOM scheme is used are how to select the methods to be combined, and how to partition the system into high- and low-level layers. These two factors are closely related. When using the ONIOM method, one must first select the high-level method. The *high*-level calculation for the *real* system is then the *target* calculation that one attempts to approach with the ONIOM method. The ONIOM results converge to the target results when the low-level method approaches the high level, or when the size of the model system approaches that of the real system. Thus, there are two ways to improve the ONIOM results, both of which should be considered.

Is there any principle for selection of the appropriate computational levels and size of the model system? It depends considerably on the properties in which one is interested and on the magnitude of error one can tolerate in the ONIOM extrapolation. Unfortunately, the studies

Table 5 IMOMO(G2MS:MP2) and experimental (in parentheses) branching ratios in Diels–Alder reactions.

Reactants	Product ratios			
	1,2, <i>cis</i>	1,2, <i>trans</i>	1,3, <i>cis</i>	1,3, <i>trans</i>
Acrylic acid + 2,4-pentadienoic acid	75.5 (61)	22.9 (22)	0.9 (9)	0.7 (8)
	<i>syn</i>	<i>anti</i>		
Ethylene + 1	99.9 (93)	0.1 (7)		
Acetylene + 1	7.2 (8)	92.8 (92)		

reported so far cover only a small number of chemical problems, so it is unlikely that the problem of interest to a specific reader would be among them. The first rule of thumb is that if one knows by intuition the major players and the minor players, the major players should be included in the model system to be treated at a high level, while the minor players should be treated at only a low level in the real system. For instance, if one is interested in chemical reactions in a solvated cluster (the real system), the reactants (and solvent molecules directly involved in the reaction) will constitute the model system. However, in some cases, one cannot identify the major players. For example, for the equilibrium structure of $\text{HCl}(\text{H}_2\text{O})_4$ modeled by Re et al.² (Figure 7), the relative energies of one neutral complex structure (a) and two ion-pair structures (b) and (c) at ONIOM[B3LYP/D95++(d,p):BLYP/D95+(d)] (HCl at high level and water molecules at low level) do not agree with either high- or low-level ordinary MO results. A detailed analysis² shows that the interaction energy in these clusters consists of the sum or difference of large two-, three-, and multibody terms, and relatively small errors introduced by ONIOM in individual terms accumulate and easily wipe out the small (~ 4 kcal/mol) difference in the total binding energy among different structures.

We have abundant experience with the study of bond dissociation, for which it is clear which ONIOM schemes to use³ [54–56]. When an appropriate combination is not available from previous studies, chemical intuition, combined with trial and error, can be used to obtain schemes that agree with experimental data. A good example of such a procedure is given for the bond energies in C_{60} [55]. A problem with this strategy is that the low-level method in ONIOM calculations often plays a role that is very different than if it were a stand-alone method, and it can therefore be difficult to use intuition to find the appropriate method. Furthermore, the number of possibilities (based on computational levels and partitioning) increases exponentially when the system becomes larger and an integrated method involving three or more layers is employed. A more direct way of finding an appropriate ONIOM method is preferred.

A systematic way of finding a correct ONIOM combination is by virtue of the substituent-value (S-value) test, which we have used in several of our recent ONIOM studies. Take, for instance, a relative energy ΔE , such as bond energy or activation energy. The S-value for a certain level is defined as the difference in ΔE between the real and the model system:

² S. Re, Y. Osamura, and K. Morokuma, work in preparation.

³ T. Vreven and K. Morokuma, work in preparation.

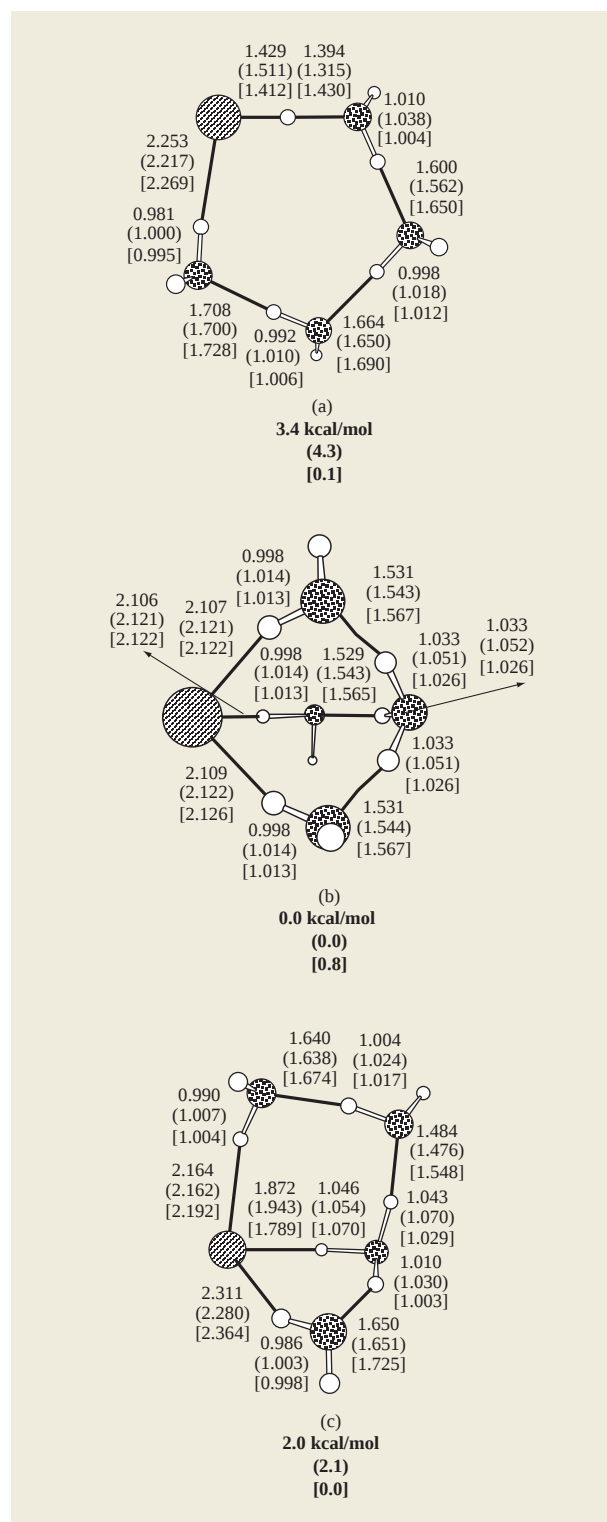
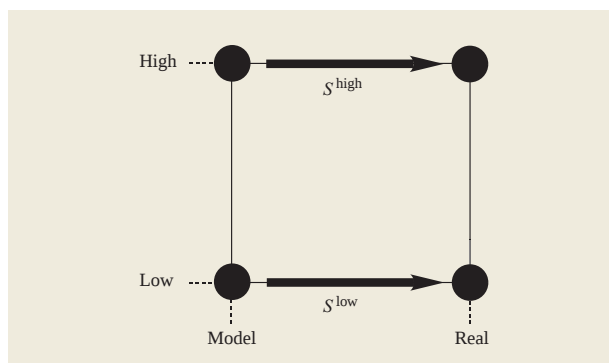


Figure 7

Optimized structures (in Å) and relative energies (in kcal/mol) of three forms of $\text{HCl}(\text{H}_2\text{O})_4$ at levels B3LYP/D95++(d,p), BLYP/D95+(d) (in parentheses), and ONIOM(B3LYP/D95++(d,p):BLYP/D95+(d)) [in brackets; the model system is $\text{HCl}(\text{H}_2\text{O})$].



Scheme 7

Definition of the S-values in the ONIOM scheme.

$$\Delta S^{\text{level}} = \Delta E_{\text{real}}^{\text{level}} - \Delta E_{\text{model}}^{\text{level}}, \quad (9)$$

i.e., the “substituent effect” on ΔE at a given level, as shown in **Scheme 7**. Using the ΔS values, the error D of the ONIOM extrapolation ΔE^{ONIOM} , compared to the target calculation $\Delta E_{\text{real}}^{\text{high}}$, can be written as

$$\begin{aligned} \Delta D &= \Delta E_{\text{real}}^{\text{high}} - \Delta E^{\text{ONIOM}} \\ &= (\Delta E_{\text{real}}^{\text{high}} - \Delta E_{\text{model}}^{\text{high}}) - (\Delta E_{\text{real}}^{\text{low}} - \Delta E_{\text{real}}^{\text{low}}) \\ &= \Delta S^{\text{high}} - \Delta S^{\text{low}}. \end{aligned} \quad (10)$$

If the ΔS value at the low level is the same as the ΔS value at the high level, the ONIOM method exactly reproduces the target method.

How can the S-value test be used in practice?

Generally, one attempts to use the ONIOM method to reproduce a target level of calculations for a series of compounds. Obtaining the ΔS^{high} value is equivalent to performing the high-level calculation for the real system one was trying to avoid. The idea of the S-value test is to calculate ΔS values at various levels, including the level one wants to use as the high level, for a test set comprising a number of the smaller compounds. These ΔS^{high} values can then be compared to the ΔS values for a variety of potential low levels, and the method closest to the ΔS^{target} values will yield the most accurate ONIOM results. Thus, using a small test set, the ONIOM combination can be calibrated and subsequently used to investigate the systems of interest. Of course, the S-value test is possible only when calculations of a representative set can be performed at the target level.

Let us consider as an example the hydrogen atom dissociation from *iso*-butane and toluene [56]:

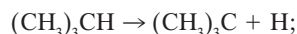
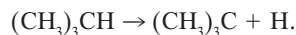


Table 6 ΔS values (kcal/mol) for hydrogen atom dissociation for *iso*-butane and toluene, employing a minimal (methane) model and B3LYP/6-31G geometries.

Level	<i>Iso</i> -butane	Toluene
G2MS(R)	−7.75(0.00)	−16.21(0.00)
UHF	−7.29(+0.46)	−26.12(−9.91)
RHF	−6.73(+1.02)	−10.28(+5.93)
UMP2	−7.72(+0.03)	+6.93(+23.14)
RMP2	−8.10(−0.35)	−15.04(+1.17)



The geometries are obtained at the B3LYP/6-31G level, and the minimal model system (methane) has been employed. In **Table 6** we show the ΔS values for the two dissociation reactions, obtained with a variety of methods [all with the 6-31G(d) basis set], together with the target level G2MS(R). The ΔS values of *iso*-butane are all quite close to the target value of −7.75 kcal/mol, and from this we would conclude that we could simply choose the cheapest of the four methods as the low level in ONIOM2 calculations. However, when we look at the ΔS value for toluene, a very different picture appears. Only the restricted MP2 method gives a ΔS value close enough to the target value, and it is clear that the other methods cannot be used low-level to study dissociation energies of compounds including aromatic substituents. This example shows the importance of a representative set to perform the S-value test. Because of the importance of testing ONIOM combinations, future versions of the ONIOM program in the Gaussian package will be able to calculate the “full square” of energies, and report the resulting S-values.

Steric effects in the reactions of transition-metal complexes

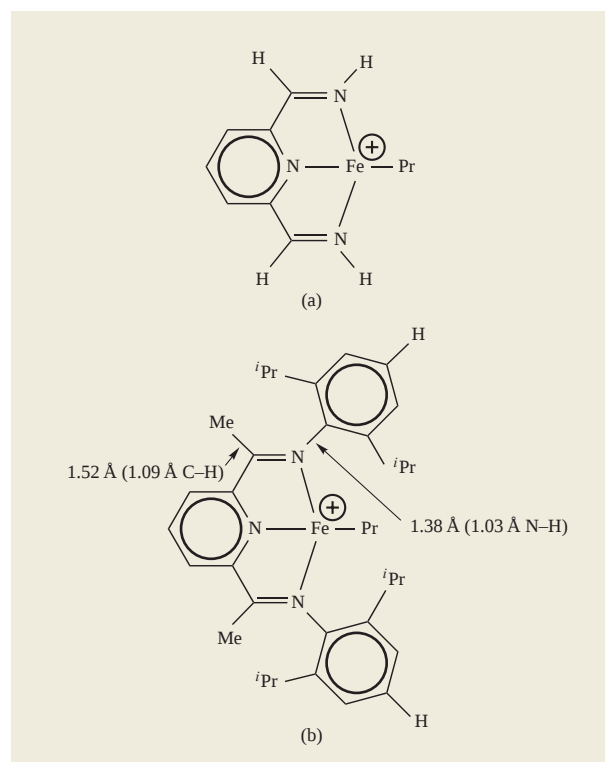
Most of the transition-metal catalysts used in homogeneous catalysis have large, often bulky substituents on the ligands. Some of them are just for synthetic convenience, but often they subtly control the reactivities of the catalysts, giving totally different products or different regio- and stereo-selectivities. The ONIOM method is ideally suited for studies of these delicate changes, and it has been used extensively for this purpose; readers are referred to recent reviews for such applications [61, 62]. Here, as an example, we show a case of olefin polymerization by a series of Fe complexes, in which the steric repulsion between bulky ligands was found to affect the electronic structure of intermediates and transition states, thereby changing the course of the reaction.

Using the IMOMM method, Khoroshun et al. [63] studied the mechanism of chain propagation and β -hydride transfer (BHT) chain termination stages of polymerization and oligomerization of ethylene by catalysts of the general formula $[2,6-(\text{CR}^1=\text{N}((2-\text{R}^2)(4-\text{R}^4)(6-\text{R}^3)\text{C}_6\text{H}_2)_2\text{C}_5\text{H}_3\text{N})\text{FeCl}_2]$. Two models of the active catalyst have been adopted, as shown in **Scheme 8**: a model “low-steric-bulk” (LSB) system and a “high-steric-bulk” (HSB) system, the actual complex studied experimentally and used to determine the effects of the bulky ligands with the IMOMM method. It was found that two axial ligands are required in order for the d_{z^2} orbital (with the tri-chelating ligand defining the equatorial xy plane) to be destabilized and for singlet to be the ground state, and that this is realized in BHT chain-termination-related species. In contrast, in the chain-propagation region of the potential energy surface (PES), only one axial ligand is present, in which, consequently, the d_{z^2} orbital is singly occupied and the singlet becomes a low-lying excited state. The calculations on the LSB system, as shown in **Figure 8**, place the lowest (singlet) BHT transition state 5.7 kcal/mol lower than the lowest (quintet and singlet) chain-propagation transition states. The inclusion of both zero-point energy and entropy corrections (namely the Gibbs free energy) notably favors higher spin states. This effect should be of the general character for highly coordinated open-shell transition-metal complexes. On the Gibbs free-energy surface of the LSB system, the lowest singlet BHT transition state is only 1.0 kcal/mol lower than the lowest quintet chain-propagation transition state. On the other hand, in the HSB system, the axial positions are sterically destabilized. The main effect of increasing the steric bulk in the axial position is the differentiation of the two ways of “saturating” the d_{z^2} orbital, one by destabilizing it, as in singlet species, and the other by populating it with the Fe d electron, in favor of the latter. On the potential surface of the HSB system, as shown in **Figure 9**, the lowest BHT transition state lies 17.6 kcal/mol higher than the lowest chain-propagation transition state. This clearly explains the experimentally observed suppression of BHT chain termination upon increase in steric bulk, and at the same time shows that different spin states dominate different pathways of the reaction of this catalytic system.

Excited states

Although most of the IMOMO studies so far have dealt with the ground state, the method is also applicable to the investigation of excited electronic states. For a two-layer ONIOM method, the energy of the excited state can be written as

$$E^{\text{ES,ONIOM}} = E_{\text{real}}^{\text{ES,low}} - E_{\text{model}}^{\text{ES,low}} + E_{\text{model}}^{\text{ES,high}} \quad (11)$$



Scheme 8

Two models used in the study: (a) Model 1: Low steric bulk (LSB): B3LYP/BS III // B3LYP/BS I; (b) Model 2: High steric bulk (HSB) IMOMM (B3LYP/BS III : MM3) // IMOMM (B3LYP/BS II : MM3).

We can distinguish two cases in the study of excited states using the ONIOM method. First, when the excitation is sufficiently localized in the model system, the effect of the low-level region, $E_{\text{real}}^{\text{ES,low}} - E_{\text{model}}^{\text{ES,low}}$, may be calculated using the ground-state energies. This is convenient when the low-level method cannot be used for excited states, for example with molecular mechanics methods. Therefore, an electronic excitation in an IMOMM (or QM/MM) calculation always reduces to the excitation in the QM region alone. When the excitation is not localized in the model system, we need to employ an IMOMO scheme with excited-state methods for both the low level and the high level. Now special care must be taken to ensure that the results correspond to the correct electronic configuration; all three subcalculations must represent the same excitation.

We recently applied IMOMO to the S_1 photoisomerization of the retinal protonated Schiff base (PSB) [64]. As a target we used an 11 non-H model for retinal PSB, calculated at the CASSCF(10e,10o)/6-31G(d). This is the system and the method used by Robb and Olivucci to determine an isomerization reaction path on

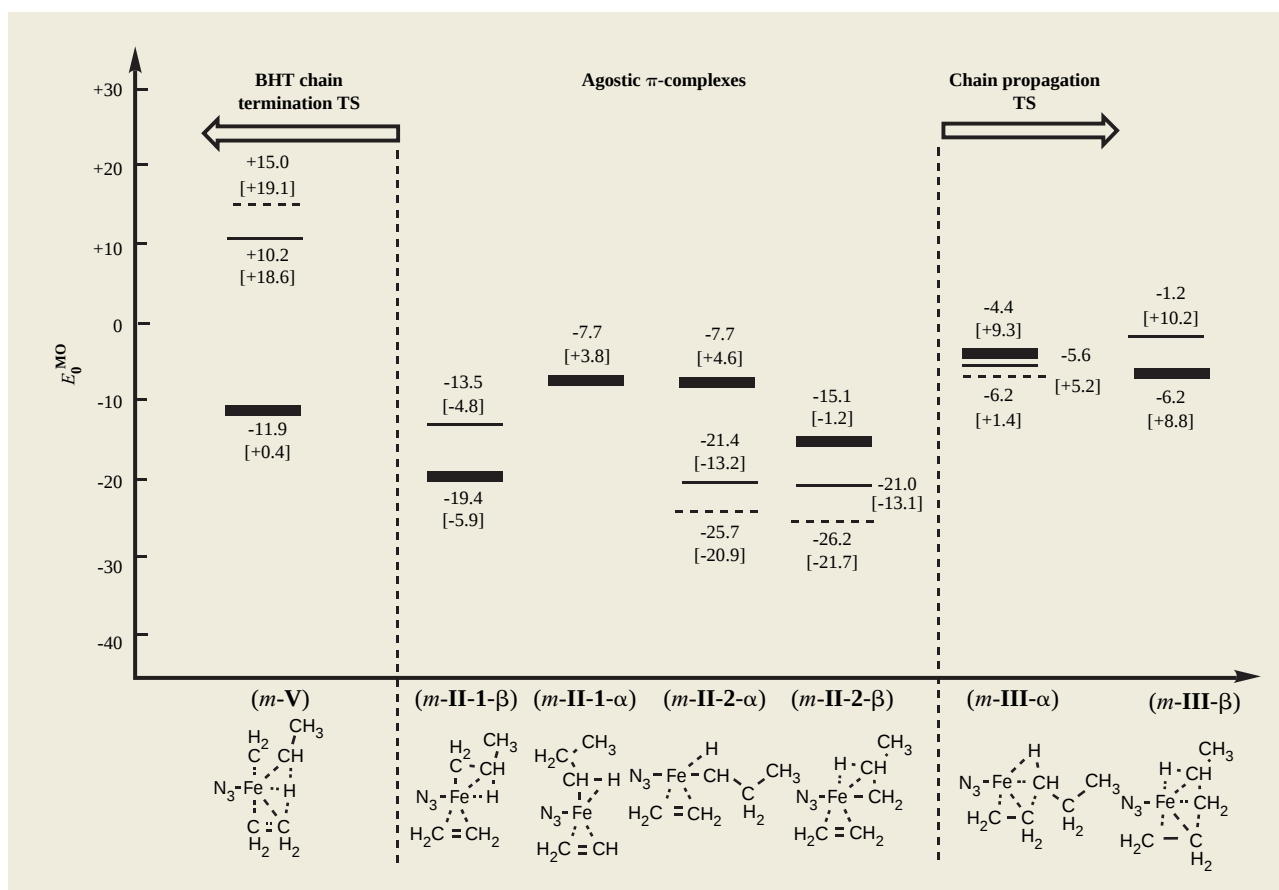


Figure 8

Competition between chain propagation and BHT chain-termination mechanism in LSB system. B3LYP/BSIII//B3LYP/BSI potential energy surfaces (in kcal/mol; Gibbs free-energy energies in square brackets) for alkyl ethylene π -complexes (*m-II*), propagation transition states (*m-III*), and BHT transition states (*m-V*). Relative to (*m-I-1-β S*) + C_2H_4 . Singlet species: thick line; triplet species: thin line; quintet species: dashed line.

the S_1 potential energy surface [65], which we attempted to reproduce using IMOMO. We tested several different partitionings and a large number of methods for the low level, both including and excluding excited-state methods. Surprisingly, the IMOMO(CASSCF^{S1}(8e,8o)/6-31G(d): UHF^{T1}/3-21G(d)) combination, with an 8 non-H model, yielded the best results. This is because the T_1 state has essentially the same electronic configuration as the S_1 state. Of course, as a stand-alone method UHF^{T1} would not reproduce the target well, but as a low-level method it reproduces the effects of the low-level region on CASSCF quite well. In Figure 10 we show the energies of both the target and the ONIOM calculations along the minimum-energy path (MEP) determined by CASSCF(10e,10o)/6-31G(d), starting at zero degrees and ending at a 68-degree torsion angle. We find that the IMOMO error is clearly very small. One could argue that the model system is not significantly smaller (3 non-H) than the real system, but

the present ONIOM calculation costs only 10% of the CPU time required for CASSCF(10e,10o). More significantly, we can easily extend our systems to the full retinal PSB and full bacteriorhodopsin.

4. The next step—Application of the ONIOM method to biological problems

The ONIOM method developed above appears to be ideally suited for the study of metalloenzyme systems as well as large molecular structures (nanostructures, fullerenes) and realistic catalytic systems (homogenous, heterogeneous). At present, we do not yet have results for such applications, although work on them is in progress.

In computationally studying large realistic models of metalloenzyme systems, one obviously wishes to maintain accuracy equivalent to or even better than that adopted for the small model systems—a good DFT method or, better, a coupled cluster method with a large basis set.

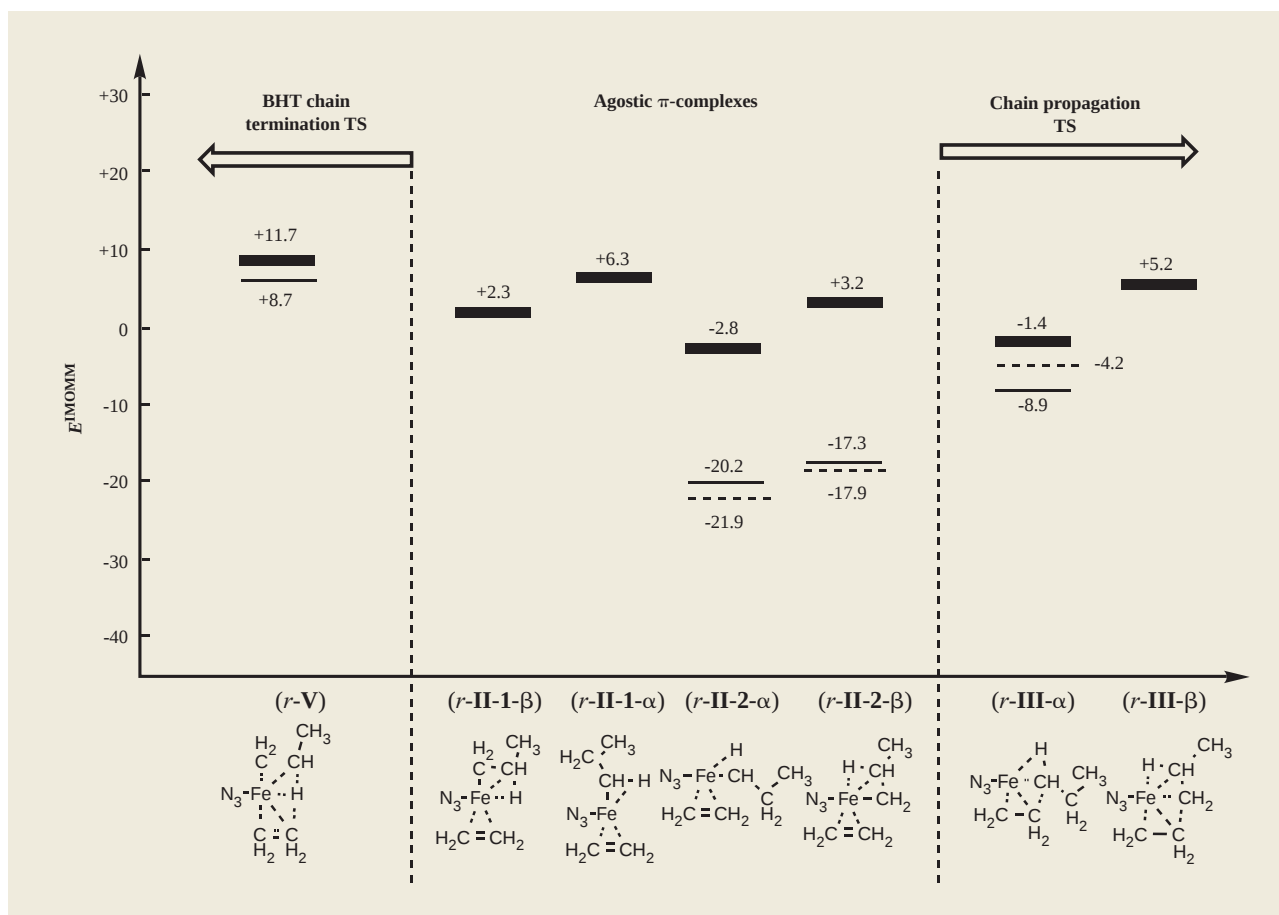


Figure 9

Competition between chain propagation and BHT chain-termination mechanism in HSB system. IMOMM(B3LYP/BSIII:MM3)/IMOMM(B3LYP/BSII:MM3) potential energy surfaces (in kcal/mol) for alkyl ethylene π -complexes ($r-II$), propagation transition states ($r-III$), and BHT transition states ($r-V$). Relative to $(r-I-1-\beta S) + C_2H_4$. Singlet species: thick line; triplet species: thin line; quintet species: dashed line.

In expanding a model, one would like to include in the calculation, in addition to the metal, the substrate, and the ligands directly coordinated with the metal (in general, protein residues and some small molecules such as water molecules), some other ligands and molecules in the neighborhood. These secondary ligands/molecules not directly involved in the reaction may be treated by an MO method of lower accuracy, such as a semiempirical MO method. In addition, the remainder of the enzyme protein will exert long-range electrostatic interaction as well as imposing steric requirements due to the fact that protein backbone structure may restrict the motion of major players such as metal, substrate, and first-shell ligands. This effect can probably be taken into account by including the remainder (or a part) of the enzyme with a molecular mechanics method. Thus, a three-layered ONIOM(MO:MO:MM) method provides an interesting

option for metalloenzyme systems. With respect to the mechanism of photoisomerization of retinal PSB in bacteriorhodopsin briefly discussed above, the ONIOM(CAS:HF:MM) combination, with the CAS:HF part for the retinal fragment and some of the nearest protein residues, augmented with a molecular mechanics method for the protein environment, appears to be appealing.

Here we wish to comment on our preference for the three-layered ONIOM(MO:MO:MM) method over the two-layered ONIOM(MO:MM) method or the QM/MM method. As mentioned briefly, proper handling of the "polarization" of the quantum-mechanical part by the electrostatic potential due to the MM part is always problematic. Often the electrostatic interaction between charges is scaled so as not to overestimate such an interaction; there is a substantial arbitrariness in this

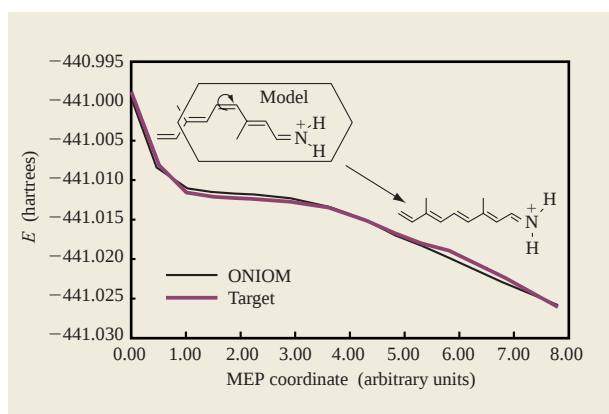


Figure 10

Target CASSCF^{S1}(10e,10o)/6-31G(d) and ONIOM(CASSCF^{S1}(8e,8o)/6-31G(d):UHF^{T1}/3-21G) curves for an S₁ photoisomerization path of an 11 non-H model for retinal protonated Schiff base.

scaling, and the QM/MM results are very much influenced by the way in which electrostatic interaction between the two parts is handled. In the three-layer ONIOM method, one introduces the second quantum-mechanical layer between the action part and the MM part. Quantum-mechanical treatment guarantees the proper polarization of the mid-layer, with no concern about scaling. Since the MM layer is far away from the first layer (the action part), the effect of the electrostatic interaction from the MM layer on the reaction in this layer is small, and the scaling becomes irrelevant.

5. Conclusions and future developments

The ONIOM method can combine different levels of MO methods with an MM method into a single integrated calculation of energy or other electronic properties, a unique feature not available in other hybrid methods. The method elegantly takes into account both electronic and steric effects of the environment or substituents on the energy, geometry, and other properties of interest, and can be applied to excited states as well as the ground state. The method is so flexible that the final choice of method and model combinations is left entirely to the user, who can tune it to be within the tolerated error (compared to either target calculations or experimental data) for the property under investigation. Generally speaking, if the target level is CCSD(T) or G2, the best choice of low level is MP2. If MP2 or DFT is the target level, HF or eventually semiempirical MO methods are good choices of low level.

Of course, ONIOM is strongly dependent on the user's selecting the correct combination of partitioning and method; otherwise, the results can be completely wrong.

However, in virtually all of the cases we eventually found an ONIOM combination that yields acceptable results. This is expected, since ONIOM should converge on the "target" with increasing size of the model system or improvement of the method used for the low level. For only those cases in which there is no "main player" for the properties of interest in the system, as in the example of the relative stability of HCl(H₂O)₄ complexes, the ONIOM method may not be suitable. One can easily verify the reliability of a prospective combination by performing the S-value test.

Possible future developments include the incorporation of the ONIOM method into the polarized continuum model for solvation [66]. A particularly interesting application of ONIOM+PCM would be to explicitly include several solvent molecules at the low level as a buffer between the high-level layer and the continuum. The effects of localized interaction of the solute with the first solvation shell could then be explicitly considered, and the effects of the continuum on the solute would be less prone to errors.

We have so far only explored potential energy surfaces in a static manner, but ONIOM could be used for other types of investigation as well. For example, we could perform direct dynamics calculations for large molecular systems using ONIOM energies and geometrical derivatives [67]. ONIOM energies and gradients can be obtained at much less cost than standard MO calculations with the same accuracy, and can be used to follow reaction dynamics of complicated molecular systems. In fact, any type of investigation of potential energy surfaces that can be carried out with conventional methods can be used with ONIOM as well.

Even with the development of ONIOM or other new tools for the study of large molecular systems including metalloenzymes, the demand for computational power for modeling biological systems using realistic models is extremely high. We depend very heavily on the development of faster and larger computers, such as Deep Blue or ASCI White built by IBM. On a smaller scale, with a grant from the National Science Foundation (CHE-0079627), we are planning to dramatically improve the computing power of the Emerson Center at Emory University so that we can actually perform such calculations.

Acknowledgment

The research presented here is partially supported by National Science Foundation Grant No. CHE-9627775. We wish to acknowledge all collaborators, visiting fellows, postdoctoral fellows, and graduate students who participated in the MMO/RNR project and the ONIOM project.

References

- (a) B. J. Waller and J. D. Lipscomb, *Chem. Rev.* **96**, 2625 (1996) and references therein; (b) E. Solomon, T. C. Brunold, M. I. Davis, J. N. Kemsley, S.-K. Lee, N. Lehnert, F. Neese, A. J. Skulan, Y.-S. Yang, and J. Zhou, *Chem. Rev.* **100**, 235 (2000) and references therein.
- R. E. Stenkamp, *Chem. Rev.* **94**, 715 (1994).
- (a) A. L. Feig and S. J. Lippard, *Chem. Rev.* **94**, 759 (1994); (b) K. E. Liu and S. J. Lippard, *Adv. Inorg. Chem.* **42**, 263 (1995); (c) B.-M. Sjöberg, *Structure of Ribonucleotide Reductase from Escherichia Coli*, Vol. 9, F. Eckstein and D. M. J. Lilley, Eds., Springer-Verlag, Berlin, 1995, pp. 192–221.
- (a) See Reference 1 and references cited therein; (b) L. Shu, J. C. Nesheim, K. Kauffmann, E. Munck, J. D. Lipscomb, and L. Que, Jr., *Science* **275**, 515 (1997).
- B. G. Fox, J. Shanklin, C. Somerville, and E. Munck, *Proc. Natl. Acad. Sci. USA* **90**, 2486 (1993).
- I. Nordlund, J. Powlowski, and V. Shingler, *J. Bacteriol.* **172**, 6826 (1990).
- J. Shanklin, E. Whittle, and B. G. Fox, *Biochemistry* **33**, 12787 (1994).
- L. M. Newman and L. P. Wackett, *Biochemistry* **34**, 14066 (1995).
- J. D. Pikus, J. M. Studts, C. Achim, K. E. Kauffmann, E. Munck, R. J. Steffan, K. McClay, and B. G. Fox, *Biochemistry* **35**, 9106 (1996).
- (a) T. Klabunde, N. Strater, R. Frohlich, H. Witzel, and B. Krebs, *J. Biol. Chem.* **259**, 737 (1996); (b) T. Klabunde and B. Krebs, *Struct. Bonding* **89**, 177 (1997).
- (a) F. Bonomi, D. M. Kurtz, Jr., and X. Cui, *J. Bioinorg. Chem.* **1**, 67 (1996); (b) E. D. Coulter, N. V. Shenvi, and D. M. Kurtz, Jr., *Biochem. Biophys. Res. Commun.* **255**, 317 (1999); (c) F. deMare, D. M. Kurtz, Jr., and P. Nordlund, *Nat. Struct. Biol.* **3**, 539 (1996).
- (a) H. Takagi, D. S. Shi, Y. Ha, N. M. Allewell, and E. C. Thiel, *J. Biol. Chem.* **273**, 18685 (1998); (b) P. Moenne-Loccoz, C. Krebs, K. Herlihy, D. E. Edmondson, E. C. Thiel, B. H. Huynh, and T. M. Loehr, *Biochemistry* **38**, 5290 (1999); (c) Y. Ha, D. S. Shi, G. W. Small, E. C. Thiel, and N. M. Allewell, *J. Bioinorg. Chem.* **4**, 243 (1999); (d) Z. Gdaniec, G. H. Sierzputowska, and E. C. Thiel, *Biochemistry* **38**, 5676 (1999); (e) J. Hwang, C. Krebs, B. H. Huynh, D. E. Edmondson, E. C. Thiel, and J. E. Penner-Hahn, *Science* **287**, 122 (2000).
- (a) V. J. DeRose, K. E. Liu, D. M. Kurtz, Jr., B. M. Hoffman, and S. J. Lippard, *J. Amer. Chem. Soc.* **115**, 6440 (1993); (b) B. G. Fox, M. P. Hendrich, K. K. Surerus, K. K. Andersson, W. A. Froland, and J. D. Lipscomb, *J. Amer. Chem. Soc.* **115**, 3688 (1993); (c) H. Thomann, M. Bernardo, J. M. McCormick, S. Pulver, K. K. Andersson, J. D. Lipscomb, and E. I. Solomon, *J. Amer. Chem. Soc.* **115**, 8881 (1993).
- (a) See Reference 1 and references cited therein; (b) J. Stubbe, *J. Biol. Chem.* **265**, 5329 (1990); (b) B.-M. Sjöberg and A. Graslund, in *Advances in Inorganic Biochemistry*, E. C. Thiel, G. L. Eichorn, and L. G. Marzilli, Eds., Elsevier, New York, 1983, ch. 5, p. 87; (c) P. Reichard and A. Ehrenberg, *Science* **221**, 514 (1983).
- (a) See Reference 1 and references cited therein; (b) P. Reinhard, *Science* **260**, 1773 (1993); (c) J. Stubbe and W. A. van der Donk, *Chem. Rev.* **98**, 705 (1998); (d) E. Mulliez and M. Fontecave, *Coord. Chem. Rev.* **185–186**, 775 (1999); (e) P. Nordlund, B. M. Sjöberg, and H. Eklund, *Nature* **45**, 593 (1990).
- W. C. Voegtli, N. Khidekel, J. Baldwin, B. A. Ley, J. M. Bollinger, Jr., and A. C. Rosenzweig, *J. Amer. Chem. Soc.* **122**, 3235 (2000) and references therein.
- M. E. Andersson, M. Hogbom, A. Rinaldo-Matthis, K. K. Andersson, B.-M. Sjöberg, and P. Nordlund, *J. Amer. Chem. Soc.* **121**, 2346 (1999).
- (a) See Reference 1 and references cited therein; (b) A. C. Rosenzweig, C. A. Frederick, S. J. Lippard, and P. Nordlund, *Nature* **366**, 537 (1993); (c) A. C. Rosenzweig, P. Nordlund, P. M. Takahara, C. A. Frederick, and S. J. Lippard, *Chem. Biol.* **2**, 409 (1995).
- P. Nordlund and H. Eklund, *J. Mol. Biol.* **232**, 123 (1993).
- D. T. Logan, X.-D. Su, A. Åberg, K. Regnström, J. Hajdu, H. Eklund, and P. Nordlund, *Structure* **4**, 1053 (1996).
- S. C. Pulver, W. A. Froland, J. D. Lipscomb, and E. I. Solomon, *J. Amer. Chem. Soc.* **119**, 387 (1997).
- S. C. Pulver, W. H. Tong, M. J. Bollinger, Jr., J. Stubbe, and E. I. Solomon, *J. Amer. Chem. Soc.* **117**, 12664 (1995).
- M. Atta, P. Nordlund, A. Åberg, H. Eklund, and M. Fontecve, *J. Biol. Chem.* **267**, 20682 (1992).
- (a) K. E. Liu, D. Wang, B. H. Huynh, D. E. Edmondson, A. Salifoglou, and S. J. Lippard, *J. Amer. Chem. Soc.* **116**, 7465 (1995); (b) K. E. Liu, A. M. Valentine, D. Wang, B. H. Huynh, D. E. Edmondson, A. Salifoglou, and S. J. Lippard, *J. Amer. Chem. Soc.* **117**, 10174 (1995); (c) K. E. Liu, A. M. Valentine, D. Qiu, D. E. Edmondson, E. H. Appelman, T. G. Spiro, and S. J. Lippard, *J. Amer. Chem. Soc.* **117**, 4997 (1995).
- E. C. Wilkinson, Y. Dong, Y. Zang, H. Fujii, R. Fraczekiewicz, G. Fraczekiewicz, R. S. Czernuszewicz, and L. Qui, Jr., *J. Amer. Chem. Soc.* **120**, 955 (1998).
- S. K. Lee, J. C. Nesheim, and J. D. Lipscomb, *J. Biol. Chem.* **268**, 21569 (1993).
- A. M. Valentine, S. S. Stahl, and S. J. Lippard, *J. Amer. Chem. Soc.* **121**, 3876 (1999).
- H. Basch, K. Mogi, D. G. Musaev, and K. Morokuma, *J. Amer. Chem. Soc.* **121**, 7249 (1999).
- M. Torrent, D. G. Musaev, and K. Morokuma, *J. Phys. Chem. B* **105**, 322 (2001).
- (a) B. D. Dunietz, M. D. Beachy, Y. Cao, D. A. Whittington, S. J. Lippard, and R. A. Friesner, *J. Amer. Chem. Soc.* **122**, 2828 (2000); (b) P. E. M. Siegbahn, *Inorg. Chem.* **38**, 2880 (1999); (c) P. E. M. Siegbahn, R. H. Crabtree, and P. Nordlund, *J. Biol. Inorg. Chem.* **3**, 314 (1998); (d) P. E. M. Siegbahn and R. H. Crabtree, *J. Amer. Chem. Soc.* **119**, 3103 (1997); (e) K. Yoshizawa, T. Ohta, and T. Yamabe, *Bull. Chem. Soc. Jpn.* **71**, 80862 (1998); (f) K. Yoshizawa, *J. Biol. Inorg. Chem.* **3**, 318 (1998); (g) K. Yoshizawa, Y. Shiota, and T. Yamabe, *Chem. Eur. J.* **3**, 1160 (1997); (h) K. Yoshizawa, Y. Shiota, and T. Yamabe, *J. Amer. Chem. Soc.* **120**, 564 (1998); (i) K. Yoshizawa, Y. Shiota, and T. Yamabe, *Organometallics* **17**, 2825 (1998) and references cited therein.
- J. B. Lynch, C. Juarez-Garcia, E. Münck, and L. Que, Jr., *J. Biol. Chem.* **264**, 8091 (1989).
- T. E. Elgren, M. P. Hendrich, and L. Que, Jr., *J. Amer. Chem. Soc.* **115**, 9291 (1993).
- (a) S. C. Pulver, W. H. Tong, J. M. Bollinger, J. Stubbe, and E. I. Solomon, *J. Amer. Chem. Soc.* **117**, 12664 (1995); (b) Y.-S. Yang, J. A. Broadwater, S. C. Pulver, B. G. Fox, and E. I. Solomon, *J. Amer. Chem. Soc.* **121**, 2770 (1999).
- (a) X.-X. Zhang, P. Fuhrmann, and S. J. Lippard, *J. Amer. Chem. Soc.* **120**, 10260 (1998); (b) D. Lee and S. J. Lippard, *J. Amer. Chem. Soc.* **120**, 12153 (1998); (c) S. Herold and S. J. Lippard, *J. Amer. Chem. Soc.* **119**, 145 (1997); (d) D. D. LeCloux, A. M. Barrios, T. J. Mizoguchi, and S. J. Lippard, *J. Amer. Chem. Soc.* **120**, 9001 (1998).
- (a) A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993); (b) C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988).
- D. A. Whittington and S. J. Lippard, in *ACS Symposium Series*, K. O. Hodgson and E. I. Solomon, Eds., American Chemical Society, Washington, DC, 1998.
- J.-P. Willems, A. M. Valentine, R. Gurbiel, S. J. Lippard, and B. M. Hoffman, *J. Amer. Chem. Soc.* **120**, 9410 (1998).

38. (a) M. Torrent, K. Mogi, H. Basch, D. G. Musaev, and K. Morokuma, work in preparation; (b) M. Torrent, H. Basch, D. G. Musaev, and K. Morokuma, *J. Phys. Chem. B* (2001), in press.
39. S.-K. Lee and J. D. Lipscomb, *Biochemistry* **38**, 4423 (1999).
40. J. D. Lipscomb, J. C. Nesheim, S.-K. Lee, and Y. Liu, *FASEB J.* **11**, 65 (1997) and references cited therein.
41. A. M. Valentine, B. Wilkinson, K. E. Liu, S. Komar-Panicucci, N. D. Priestley, P. G. Williams, H. Morimoto, H. G. Floss, and S. J. Lippard, *J. Amer. Chem. Soc.* **119**, 1818 (1997).
42. For a recent review, see J. Gao in *Reviews in Computational Chemistry*, Vol. 7, K. B. Lipkowitz and D. B. Boyd, Eds., VCH, New York, 1996, p. 119.
43. N. L. Allinger, MM3(92); Quantum Chemistry Program Exchange, Bloomington, IN, 1992.
44. S. Humbel, S. Sieber, and K. Morokuma, *J. Chem. Phys.* **106**, 1959 (1996).
45. M. Svensson, S. Humbel, R. D. J. Froese, T. Matsubara, S. Sieber, and K. Morokuma, *J. Phys. Chem.* **100**, 19357 (1996).
46. Gaussian 98, Revision A.1, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, and J. A. Pople, Gaussian, Inc., Pittsburgh, 1998.
47. F. Maseras, *Top. Organomet. Chem.* **4**, 165 (1999).
48. R. D. J. Froese and K. Morokuma, in *The Encyclopedia of Computational Chemistry*, P. v. R. Schleyer, N. L. Allinger, T. Clark, J. Gasteiger, P. A. Kollman, H. F. Schaefer III, and P. R. Schreiner, Eds., John Wiley, Chichester, U.K., 1998, p. 1245.
49. T. Vreven and K. Morokuma, *J. Comp. Chem.*, in press.
50. S. Dapprich, I. Komáromi, K. S. Byun, K. Morokuma, and M. J. Frisch, *J. Mol. Str. (Theochem)* **461-462**, 1 (1999).
51. N. Koga and K. Morokuma, *Chem. Phys. Lett.* **172**, 243 (1990).
52. D. Bakowies and W. Thiel, *J. Phys. Chem.* **100**, 10580 (1996).
53. M. Svensson, S. Humbel, and K. Morokuma, *J. Chem. Phys.* **105**, 3654 (1996).
54. R. D. J. Froese and K. Morokuma, *J. Phys. Chem. A* **103**, 4580 (1999).
55. R. D. J. Froese and K. Morokuma, *Chem. Phys. Lett.* **305**, 419 (1999).
56. T. Vreven and K. Morokuma, *J. Chem. Phys.* **111**, 8799 (1999).
57. R. D. J. Froese, S. Humbel, M. Svensson, and K. Morokuma, *J. Phys. Chem. A* **101**, 227 (1997).
58. J. M. McBride, *Tetrahedron* **30**, 2009 (1974).
59. R. D. J. Froese, J. M. Coxon, S. C. West, and K. Morokuma, *J. Org. Chem.* **62**, 6991 (1997).
60. J. M. Coxon, R. D. J. Froese, B. Ganguly, A. P. Marchand, and K. Morokuma, *Synlett.* **11**, 1681 (1999).
61. F. Maseras, in *Topics in Organometallic Chemistry*, Vol. 4, J. M. Brown and P. Hofmann, Eds., Springer, Berlin, 1999, pp. 166-191.
62. F. Maseras, in *Computational Organometallic Chemistry*, T. Cundari, Ed., Marcel Dekker, New York, 2000, in press.
63. D. V. Khoroshun, D. G. Musaev, T. Vreven, and K. Morokuma, *Organometallics* (2001), in press.
64. T. Vreven and K. Morokuma, *J. Chem. Phys.* (2001), in press.
65. M. Garavelli, T. Vreven, P. Celani, F. Bernardi, M. A. Robb, and M. Olivucci, *J. Amer. Chem. Soc.* **120**, 1285 (1998).
66. C. Amovilli, V. Barone, R. Cammi, E. Cancès, M. Cossi, B. Mennucci, C. S. Pomelli, and J. Tomasi, *Adv. Quantum Chem.* **32**, 227 (1999).
67. W. Chen, W. L. Hase, and H. B. Schlegel, *Chem. Phys. Lett.* **228**, 436 (1994).

Received July 24, 2000; accepted for publication
September 24, 2000

Keiji Morokuma *Cherry L. Emerson Center for Scientific Computation and Department of Chemistry, Emory University, Atlanta, Georgia 30322 (morokuma@emory.edu).* Dr. Morokuma received a Ph.D. degree from Kyoto University; he did postdoctoral work at Columbia and Harvard universities and held professorships at the University of Rochester and the Institute for Molecular Science, Okazaki, Japan. Since 1993, he has been William H. Emerson Professor of Chemistry and Director, Cherry L. Emerson Center for Scientific Computation at Emory University. His research interests are in theoretical studies of intermolecular interactions, chemical reaction mechanisms and dynamics, reactions of transition-metal complexes, and enzymatic reactions of metalloenzymes, as well as the development of new theoretical methods. Group home page: <http://euch4m.chem.emory.edu/>.

Djamaladdin G. Musaev *Cherry L. Emerson Center for Scientific Computation and Department of Chemistry, Emory University, Atlanta, Georgia 30322 (dmusaev@emory.edu).* Dr. Musaev received a Ph.D. degree from the Kurnakov Institute of General and Inorganic Chemistry, Moscow. He was a Senior Researcher at the Institute of New Chemical Problems, Chernogolovka, USSR, before accepting a JSPS Postdoctoral Fellowship at the Institute for Molecular Science, Okazaki, Japan. In 1993, he joined the Emerson Center at Emory University as an Associate Scientist, and in 2000 was promoted to Manager and Principal Scientist. His research interests are in theoretical studies of reactions of transition-metal complexes and metalloenzymes.

Thom Vreven *Cherry L. Emerson Center for Scientific Computation and Department of Chemistry, Emory University, Atlanta, Georgia 30322.* Dr. Vreven received a Ph.D. degree from King's College, London; he is a Postdoctoral Fellow at Emory University, and will be joining Gaussian, Inc., in the immediate future. His specialty is the development of the ONIOM method and code, as well as its applications to various problems, including bond energies, solvation energies, and photoisomerization of bacteriorhodopsin.

Harold Basch *Cherry L. Emerson Center for Scientific Computation and Department of Chemistry, Emory University, Atlanta, Georgia 30322 (permanent address: Department of Chemistry, Bar Ilan University, Ramat Gan, Israel).* Dr. Basch received his Ph.D. degree from Columbia University; he did his postdoctoral work at Bell Laboratories, and he was Principal Scientist at Ford Motor Company before becoming a professor at Bar Ilan University, Ramat Gan, Israel. He has been a Visiting Fellow and a frequent visitor at the Emerson Center. His research interests are in theoretical studies of inorganic reactions and enzymatic reactions.

Maricel Torrent *Cherry L. Emerson Center for Scientific Computation and Department of Chemistry, Emory University, Atlanta, Georgia 30322.* Dr. Torrent received her Ph.D. from the University of Girona, Spain; she held a Spanish Government Postdoctoral Fellowship at Emory University, and continues to be a Postdoctoral Fellow there. Her research interests are in theoretical studies of reactions of metalloenzymes and transition-metal complexes.

Dmitry V. Khoroshun *Cherry L. Emerson Center for Scientific Computation and Department of Chemistry, Emory University, Atlanta, Georgia 30322.* Mr. Khoroshun graduated from the Higher Chemical College of the Russian Academy of Sciences in Moscow. He is a graduate student at Emory University doing theoretical studies of the reaction mechanisms of transition-metal complexes and other systems. In addition, he is the system manager of a 110-CPU PC cluster of the Morokuma group.